Analysis of the natural product biosynthesis in gliding bacteria using *in vitro* assays

Dissertation

zur

Erlangung des Doktorgrades (Dr.rer.nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Mahsa Mir Mohseni

aus

Teheran

Bonn, 2016

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. G. M. König
2. Gutachter: Prof. Dr. Gerd Bendas

Tag der Promotion: 16.02.2016

Erscheinungsjahr: 2017

Für meine Eltern und Sina

Acknowledgements

My PhD has been an amazing experience and this could not be done without helping and guidance of my supervisors, my Colleagues and my family.

First of all I would like to express my gratitude to my supervisor Prof. Dr. G.M. König for giving me a wonderful opportunity to be involved in her research group. I appreciate her for her immense knowledge, motivation and patience. Every result described in this thesis was not possible to obtain without her continuous guidance and encouragement. I would like to thank her for proofreading of my thesis as well as supporting me for the scholarship applications.

Many thanks to Prof. Dr. Gerd Bendas for co-examination of my thesis as well as Prof. Dr. Karl Schellander and PD Dr. Anke Schiedel for their presentation in my examination committee.

I would like to thank International Graduate School BIOTECH-PHARMA for the financial support during my PhD.

I must acknowledge Dr. T.F. Schäberle for his time and professional support. Additionally I would like to thank him for many constructive discussions and for proofreading some parts of my work.

I would also like to thank Dr. Stefan Kehraus to know the answer of all my questions about the analytical chemistry and for his assistance with GC-MS and NMR measurements.

Dr. Marc Sylvester (Institute of Biochemistry and Molecular Biology at the University of Bonn) I want to thank him for the mass spectrometric measurements under the "PPant ejection assays".

Dr. Thomas Höver for introducing this project to me at the start point of my PhD.

My special thanks goes to my present and past colleagues Paul Barac, Antonio Davila, Christian Siba and Isabella Schamari for all scientific discussions and good advice. I would like to thank you for sharing all the good and hard times with me and for all interesting trips.

I wish to thank E. Goralski for all great technical help as well as providing the friendly atmosphere in S1 labs, also for all wonderful discussions.

I would like to thank all members of AG König specially Alexander Bogdanov, Friederike Lohr, Jan Schrör, Peter Hufendiek, Raphael Reher, Fayrouz El Maddah and Mamona Nazir for all the scientific exchange and non-scientific chat.

I also would like to thank Ms Edith Neu for providing the bacterial media and Ms. Ekatarina Eguereva for LC-MS measurements.

I thank Mr. Thomas Kögler for helping me by all of my problems with computers and lab equipment.

I would like to thank Kristel Martínez and Aishwarya Ghule from the International Graduate School BIOTECH-PHARMA for all great moments we spend together.

I want to express my particular thanks from my heart to my family, my parents and my sister for their encouragement and assisting me during my whole life. My life could be dark and meaningless without them.

At the end my special gratitude goes to my spouse Dr. Sina Seifi for all his educational support and advice for writing my dissertation. It wasn’t possible to do this work without his care and motivation.

Contents

Abbreviations

aa	Amino acid
ACP	Acyl Carrier Protein
ADC	Amino-4-deoxychorismic acid
A domain	Adenylation domain
AHBA	Amino-5-hydroxybenzoate
APS	Ammoniumperoxosulfate
Asn	Asparagine
AT	Acyl Transferase
BAC	Bacterial artificial chromosome
BLAST	Basic local alignment search tool
bp	Base pair
BTP	Bis-tris propane
C domain	Condensation domain
CHC	Cyclohexanecarboxylate
CIAP	Calf intestinal alkaline phosphatase
CoA	Coenzyme A
CP	Carrier Protein
Cyc domain	Cyclisation domain
CYP450	Cytochrome P450
Da	Dalton
DAD	Diode array detector
DAHP	3-deoxy-D-arabino-heptulosonate 7-phosphate
ddATP	2', 3'-dideoxyadenisine 5'-triphosphate
ddCTP	2', 3'-dideoxycytidine 5'-triphosphate
ddGTP	2', 3'-dideoxyguanosine 5'-triphosphate
ddTTP	2', 3'-dideoxythymidine 5'-triphosphate
DH	Dehydratase
DHQ	Dehydroquinic acid

DMAP	Dimethylaminopyridine
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
Dpg	Dihydroxyphenylglycine
EBI	European bioinformatics institute
EDC	Ethyl-3-(3-dimethylaminopropyl)carbodiimide
E domain	Epimerization domain
EDTA	N,N,N´,N´-Ethylendiamintetraacetat
E4P	Erythrose 4-phosphate
ER	Enoylreductase
ESAC	*E. coli*-Streptomyces artificial chromosome
ESI	Electrospray ionization
FAD	Flavin adenine dinucleotide
F domain	Formylation domain
FMN	Flavinmononucleotid
Formyl-THF	Formyltetrahydrofolate
FPLC	Fast protein liquid chromatography
FT	Flow through
GC	Gas Chromatography
HBA	Hydroxy benzoic acid
HMG-CoA	Hydroxy-3-methylglutaryl-coenzyme A
Hpg	Hydroxyphenylglycine
HPLC	High performance liquid chromatography
Ht	Hydroxytyrosine
IPTG	Isopropyl-β-D-thiogalactosid
IRMPD	Infrared multiphoton dissociation
kb	Kilobases
kDa	Kilodalton
KR	Ketoreductase
KS	Ketosynthase

LC-MS	Liquid chromatography–mass spectrometry
LCMT	Leucine carboxyl methyltransferase
Leu	Leucine
MALDI	Matrix-assisted laser desorption/ionization
MIC	Minimum inhibitory concentration
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Mass spectrum
NADPH	Nicotinamide adenine dinucleotide phosphate
NCBI	National center for biotechnology information
NI-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OD	Optical density
O-MT	*O*-methyltransferase
OR	Oxidoreductase
ORF	Open reading frame
oriT	Origin of transfer
Ox domain	Oxidation domain
P	Pellet
PABA	*p*-Aminobenzoate
PAC	P1-derived artificial chromosome
PAGE	Polyacrylamide gel electrophoresis
PAPS	3′-phosphate 5′-phosphosulfate
PAT	Prephenate aminotransferase
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
Phd	Prephenate dehydrogenase

PK	Polyketide
PKS	Polyketide Synthase
PP	Protein phosphatase
PPant	Phosphopantetheinyl
PPTase	Phosphopantetheinyl transferase
R domain	Reduction domain
RNAP	DNA-dependent RNA polymerase
RP	Reversed-Phase
NRPS	Non-ribosomal peptide synthetase
SA	Shikimic acid
SAH	S-adenosyl-homocysteine
SAM	S-adenosylmethionine
SAP	Shrimp alkaline phosphatase
SDS	Sodium dodecylsulfate
SNAC	N-acetylcysteamine
sp.	Species
ST	Sulfotransferase
TBE	Tris/Borate/EDTA
TCA	Tricarboxylic acid
T domain	Thiolation domain
TE	Thioesterase
TEMED	Tetramethylethylendiamin
UV	Ultraviolet

1 Introduction

1.1 Natural products as a source of leads in drug discovery

This study focuses on gliding bacteria including myxobacteria cyanobacteria and *Herpetosiphon* spp. as a source of bioactive secondary metabolites.

Since the discovery of penicillin, an efficient antibacterial agent, natural products and their semisynthetic derivatives have been considered as a rich resource for drug discovery. Many microorganisms produce natural products, some of which have important functions for the producer organism, i.e. for scavenging of iron or quorum sensing (Finking & Marahiel, 2004). Others represent a promising source of potential therapeutical agents. Of all the medicines that were based on small molecules introduced between 1981 and 2010, 34% were natural products or derive from natural sources (Harvey & Harvey, 2008). They cover a range of therapeutic indications including: antibacterial (vancomycin), antifungal (amphotericin B), immunosuppressant (cyclosporin A), antitumor (bleomycin) and cholesterol-lowering agents (mevastatin) (Cragg & Newman, 2013; Mishra & Tiwari, 2011).

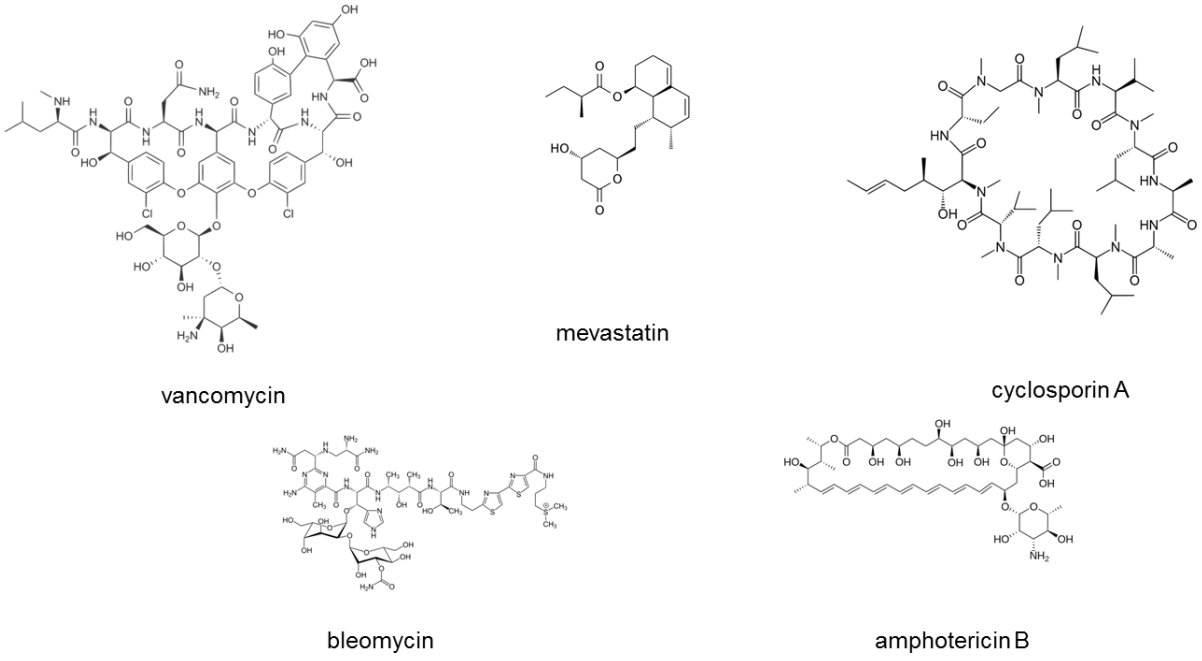


Figure 1-1. Structures of some important natural products derived from polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS) or mixed PKS/NRPS pathways.

One of the significant sources for secondary metabolite production are microorganisms. Considering the impressive number of bioactive isolated molecules, most of the bacterial sources belong to the actinomycetes. Despite their importance in generation of known secondary metabolites, actinomycetes comprise a small minority of the microbial antibiotic producers (Clardy et al., 2006). Therefore, new bacterial sources of natural products are urgently needed in order to find new structural classes. Recently using several strategies, i.e. expanded conventional culturing approaches, novel culture methods, heterologous DNA-based methods and metagenomics assist the discovery of new natural products by expanding the range of bacteria that can be tapped for drug leads (Ahmad & Aqil, 2009).

1.2 Secondary metabolites of gliding bacteria

Gliding bacteria, as found in the taxa Chloroflexi, Proteobacteria, Bacteroidetes and Cyanobacteria, are one of the example of a previously ignored group of organisms regardless of their competency for producing bioactive metabolites. Many of the members of these taxa have shown to be producing a variety of bioactive metabolites, due to rather large genome sizes (Nett & König, 2007).

Especially myxobacteria have been found to be prolific source of secondary metabolism, which places them among the well-known natural product producers, i.e. actinomycetes and *Bacillus* species. Compounds form this group of bacteria show unusual compositions with around 100 different core structures and 500 derivatives. They include antifungal (leupyrrin), antibacterial (corallopyronin), cytotoxic (tubulysin) and antiviral compounds (phenoxan) as well as many others (Weissman & Müller, 2010). Analogous to myxobacteria, cyanobacteria from different environments are another rather inaccessible source of novel chemicals (Welker et al., 2012). Several promising cyanobacterial compounds are found to target tubulin or actin filaments in eukaryotic cells (Jordan & Wilson, 1998). In this context curacin A and dolastatin, have been in preclinical and/or clinical trials as potential anticancer drugs (Taori et al., 2009; Chang et al., 2004). Cyanobacteria and myxobacteria often share a similar way for producing secondary metabolites, i.e. mixed polyketide-nonribosomal peptide system (PKS/NRPS) (Silakowski et al., 2001).

1.3 Biosynthetic machineries producing nonribosomal peptides and polyketides

Nonribosomal peptides (NRPs) and polyketides (PKs) are two prominent families of natural products which are the product of so-called nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), respectively.

NRPSs and PKSs contain numerous enzymatic domains organized into functional units termed modules. One module consists of the domains required for the incorporation of a single monomer into the product (von Döhren et al., 1997).

One of the central tenet which shared between these enzymes is utilizing a chain elongation process which proceeds via covalent tethering of all the substrates, intermediates, and products. This can be achieved by incorporation of small noncatalytic domains, called carrier proteins in the biosynthetic assembly lines. These carrier proteins can be eider the acyl carrier proteins (ACPs) in case of biosynthesis of PKs or peptidyl carrier proteins (PCPs) in biosynthesis of NRPs (Mercer & Burkart, 2007; Weissman & Müller, 2008).

For enzymatic assembly lines to become functional, every carrier proteins (also called thiolation (T) domain) should be modified posttranslationally by the addition of a phosphopantetheinyl (PPant) prosthethic arm, derived from coenzyme A. Therefore, the *apo*-ACP converts to the *holo*-ACP by transferring the PPant arm to a conserved serine residue of *apo*-ACP (Lai, Koglin, & Walsh, 2006). Activation of the CPs are catalyzed by a superfamily of enzymes called phosphopantetheinyl transferases (PPTase) (Figure 1-2).

The growing chain of the biosynthetic product tethers to the terminal sulfhydryl of a PPant by generating the thioester bond.

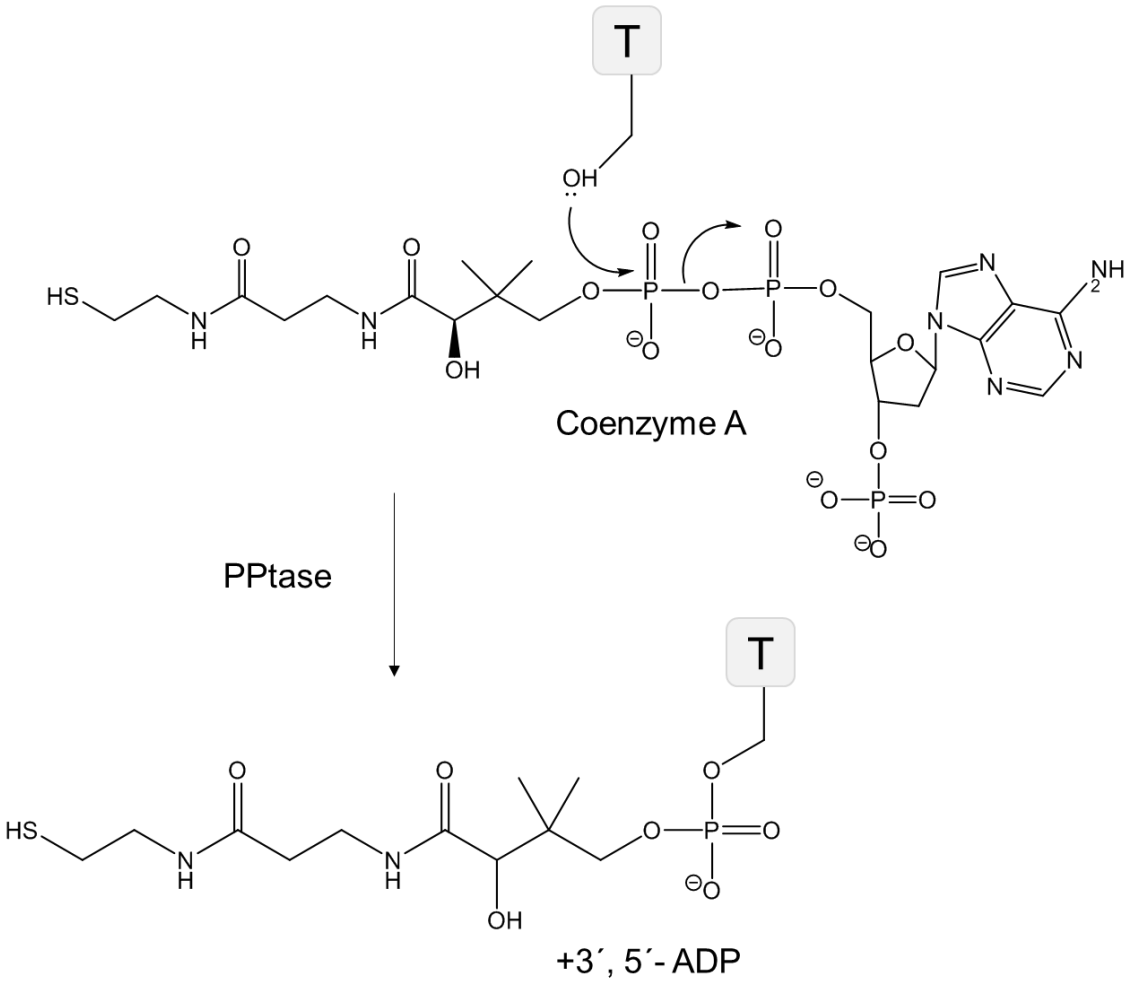


Figure 1-2. Phosphopantetheinyl transferase (PPTase) catalyzes the transfer of the 4′-phosphopantetheinyl (PPant) group to the conserved Ser residue in carrier proteins like peptidyl carrier proteins (PCP) and acyl carrier proteins (ACP). T: thiolation domain (Sattely et al., 2008).

NRPSs and PKSs use an analogous strategy for the assembly of NRPs and PKs, respectively. They both use the sequential condensation of amino acids and short carboxylic acids as well as utilizing the same 4'-phosphopantethein prosthetic group to channel the growing peptide or polyketide intermediates during the elongation

processes (B Shen et al., 2001). Structural insights into PKS and NRPS machineries will be described in the following.

1.3.1 Polyketide synthases

By exploring orcinol, the first isolated polyketide by Collie and Myers in 1893, this group of metabolite entered research. Nowadays many thousands of small molecule natural products are polyketides with complex chemical structures and enormous pharmaceutical potential. They are the products of PKS.

PKs are synthesized through decarboxylative condensation of carboxylic acids. They utilize the thioesters of monoacyl groups such as acetyl-, propionyl-, and benzoyl-CoAs, or structural variants, such as malonamyl-CoA or methoxymalonyl-CoA as chain starter units as well as malonyl-CoA and methylmalonyl-CoA as the elongation units (Fischbach & Walsh, 2006).

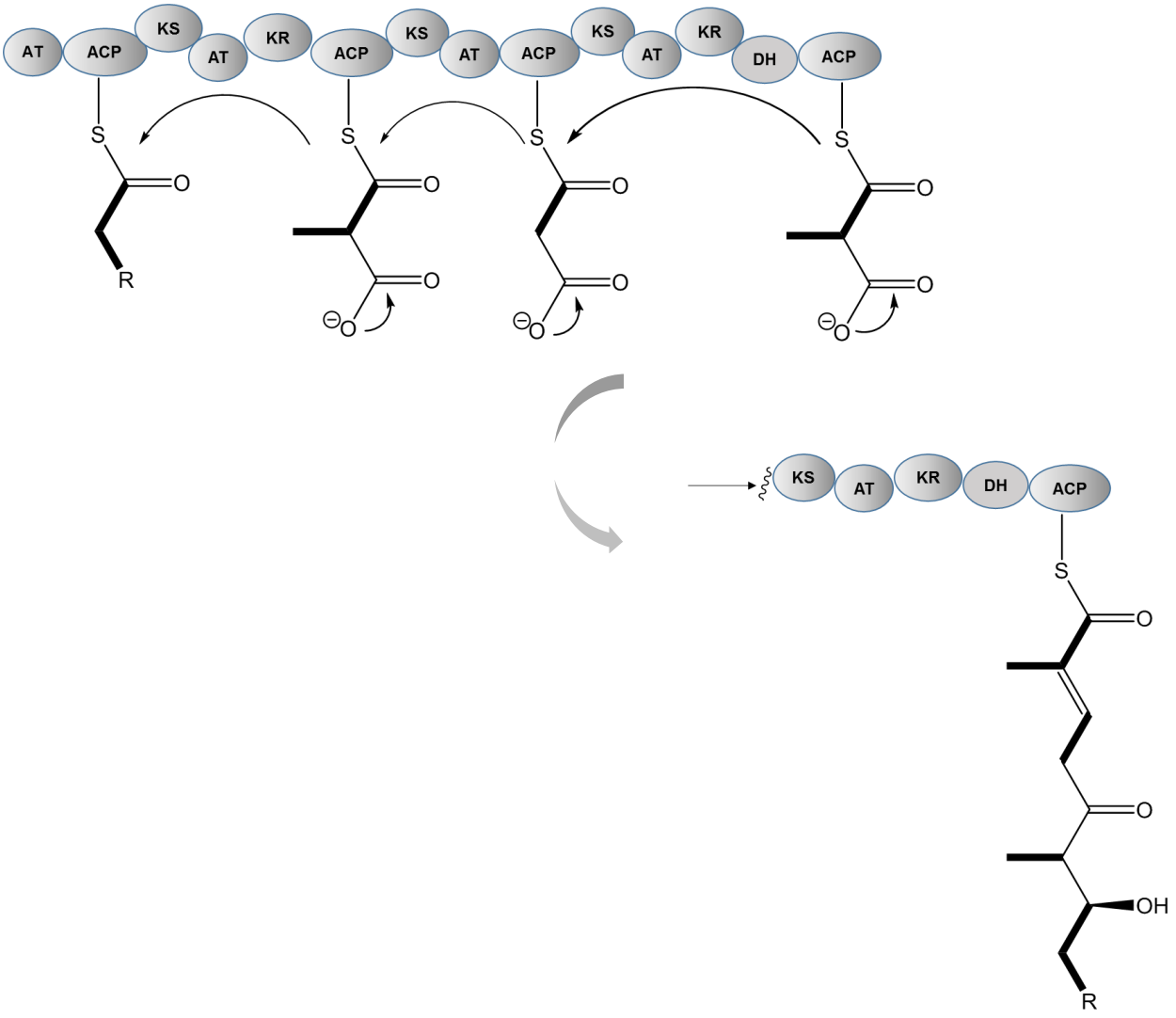
PKSs are multifunctional enzymes. Their assembly lines have a modular architecture, and modules can be further subdivided into distinct catalytic domains. The minimal PKS module consist of a ketosynthase (KS), an acyltransferase (AT) and an ACP. The product release from the PKS-template by a thioesterase (TE) domain, occurs either by hydrolysis or macrocyclization (Kopp & Marahiel, 2007).

Three types of bacterial PKSs are known to date. They are grouped according to their sequences, primary structures and catalytic mechanisms, including type I, II, and III PKS (D. Yu et al., 2012). Type I PKSs are multifunctional enzymes that are organized into modules. They harbor a set of distinct, noniteratively acting domains responsible for the catalysis of one cycle of polyketide chain elongation (Figure 1-3, A). This type of PKS can be found for the biosynthesis of macrolides, polyethers and polyenes such as erythromycin A (Weissman & Müller, 2010). For further structural diversity, PKS extension module, may be present with modifying domains, such as ketoreductase (KR), dehydratase (DH) and enoylreductase (ER). Therefore, the degree of reduction and the substitution pattern can differ for each cycle.

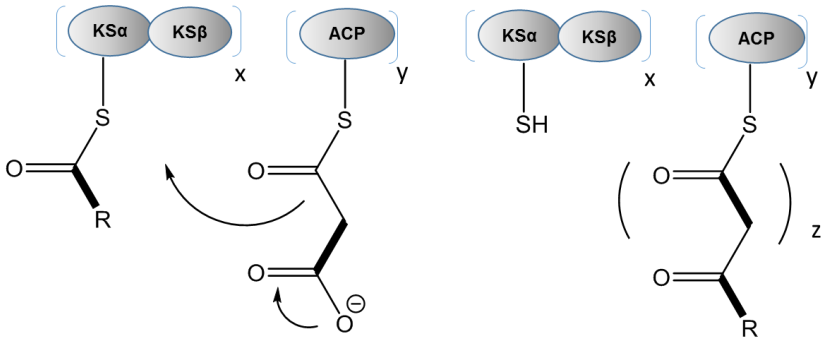
Type II PKSs are known for the biosynthesis of therapeutically important aromatic polyketides, i.e. iterative use of domains. They consist of monofunctional enzymes that work iteratively (Figure 1-3, B). Biosynthesis of actinorhodin in *Streptomyces coelicolor* A3 (2) is an example of type II PKS domains (Kim & Yi, 2012).

Type III PKSs are widespread and were initially discovered in plants and bacteria. Type III PKSs are also known as chalcone synthase. They are condensing enzymes that lack ACP and act directly on acyl CoA substrates (Figure 1-3, C). (Ben Shen, 2003; Cheng et al., 2003).

A) Type I PKS



B) Type II PKS



C) Type III PKS

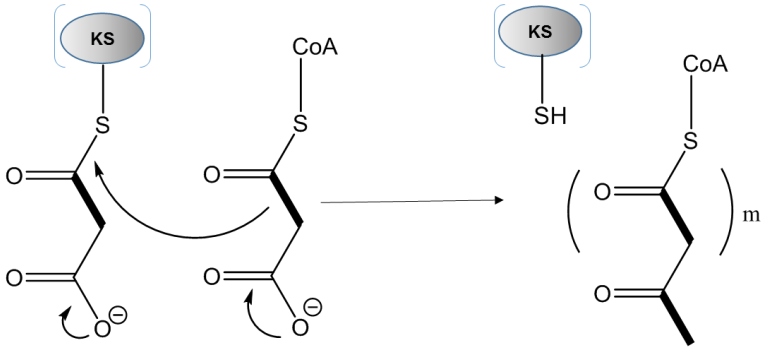


Figure 1-3. Different types of PKSs. A. Type I (noniterative); B. Type II (iterative), and C. Type III (ACP-independent and iterative) (Ben Shen, 2003).

1.3.2 Nonribosomal peptide synthetases

NRPs are a large family of structurally diverse secondary metabolites which are generated by NRPSs. NRPs represent diverse biological active compounds, including antibiotics (daptomycin), antitumor (bleomycin) as well as immunosuppressants (cyclosporin) (Strieker et al., 2010).

NRPS function almost in the same way as PKS, however they differ in their substrates and domains (Walsh, 2008). They catalyze the generation of NRPs through a thiotemplate mechanism, independent of ribosomes, from simple building blocks. Unlike the ribosomal proteins with the limitation of the obligatory usage of proteinogenic amino acids, NRPs are using of a diverse pool of building blocks. They can be composed of D- and L-amino acids, proteinogenic and non-proteinogenic amino acids, hydroxy acids, carboxylic acids like aryl acids, β-amino acids, as well as many other unusual units (Schwarzer et al., 2003). The yielded NRPs from this process can have linear, cyclic, or branched cyclic structures. They also can be modified by glycosylation, *N*-methylation, or acylation (Kleinkauf & Von Döhren, 1996; Lee et al., 2005).

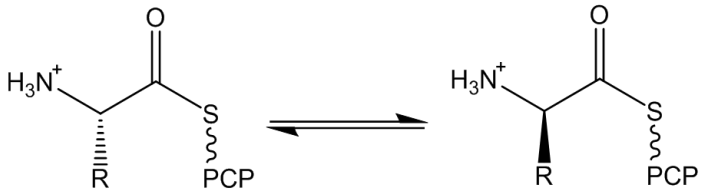
The minimal NRPS-module comprises an AMP-binding adenylation (A) domain, a condensation (C) domain and a PCP domain. However, the initiation NRPS-modules is only composed of an A-PCP didomain (Marahiel, 2009). Analogous to PKSs, synthesis of the peptide on an NRPS assembly line is terminated by a TE domain which is normally located in the termination module. This last reaction is done in two steps. First, the full length peptide, which is attached to the terminal PCP is transferred to the conserved serine of the TE domain to generate an acyl-O-TE. Afterward, this construct is cleaved off by a regio- and stereoselective intramolecular macrocyclization. However, the simple hydrolysis results in the release of a linear peptide, as observed in the formation of pyochelin and yersiniabactin (Trauger et al., 2000; Marahiel, 2009).

Apart from the canonically arranged domains many other auxiliary domains can also be integrated within an NRPS elongation module (Figure 1-4). These domains are including epimerization (E) domains which generate D-configurated amino acids from the corresponding L-isomers, *N*-methyltransferase (MT) domains, which catalyze the transferring of a methyl group to the nitrogen atoms of the peptide backbone of NRPs, formylation (F) domains responsible for the N-formylation by means of the cofactor N-formyltetrahydrofolate, and reduction (R) domains, which catalyze the reduction of C-terminal carboxy group to an aldehyde or to the corresponding alcohol by using NADPH as cofactor (Silakowski et al., 1999; Du et al., 2000).

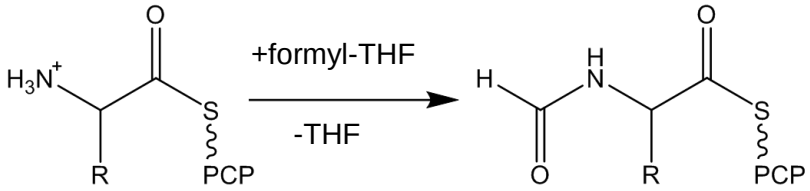
Two other frequent auxiliary domains which catalyze the occurrence of small heterocyclic rings in the biosynthesis of NRPs are cyclization (Cyc) domains and oxidation (Ox) domains.

The Cyc domain is responsible for the generation of thiazoline or oxazoline rings from incorporated threonine, serine or cysteine, respectively, e.g. generation of thiazoline rings in bacitracin A and oxazoline rings in vibriobactin. To convert thiazoline or oxazoline to the more stable forms, i.e. thiazole and oxazole rings, an additional Ox domain is required which use flavin mononucleotide (FMN) as cofactor. This can be observed in the biosynthesis of bleomycin and myxothiazol (Marahiel, 2009).

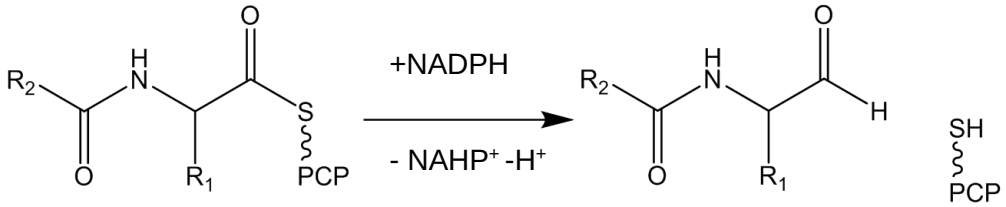
E-Domain



F-Domain



R-Domain



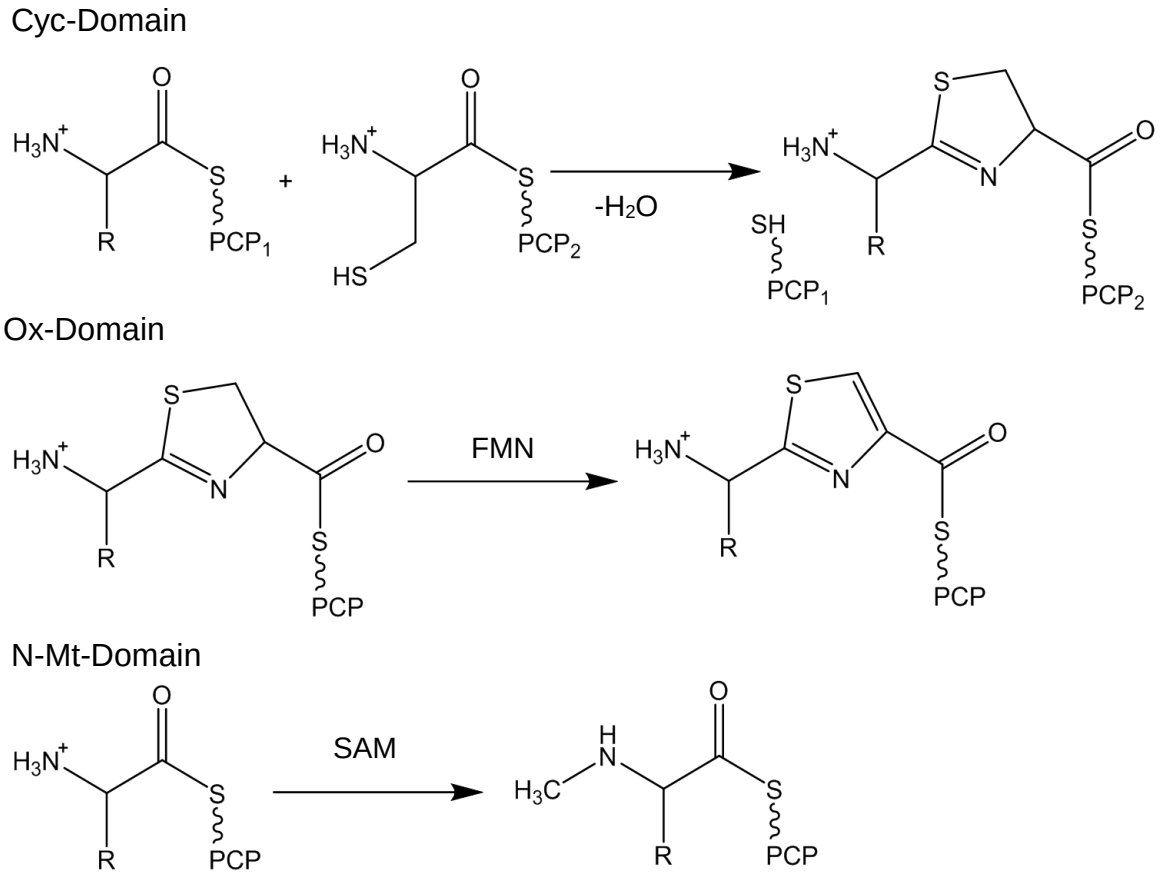


Figure 1-4. Reactions performed by NRPSs domains.

PCP, peptidyl carrier protein domain; E, epimerization domain; F, formylation domain; R, racemization domain; Cyc, cyclisation domain; Ox, oxidation; N-Mt, *N*-methyltransferase domain; FMN, Flavinmononucleotid; SAM, S-adenosyl methionine; Formyl-THF, formyltetrahydrofolate; NADPH, nicotinamide adenine dinucleotide phosphate (From Schwarzer et al., 2003).

NRPSs fall into three groups, i.e. Type A linear; Type B iterative, and Type C nonlinear. Type A NRPSs are well known for their single usage of each module during the biosynthesis and the three core domains are arranged in the order C-A–PCP. This type can be found in the biosynthesis of tyrocidine, bacitracin, surfactin and actinomycin (Konz et al.,1997; Cosmina et al., 1993).

Iterative NRPSs (Type B) use all of their modules or domains more than once to produce multimeric compounds from repeated smaller peptide sequences. Therefore, the monomer chains are transferred onto the active site serine residue of the C terminal TE domain, thereby deacylating the last PCP and thus regenerating the NRPS for the next chain assembly. Eventually oligomerization of the yielded monomers takes place on the TE domain. Then the final product will be released from the assembly line. Biosynthesis of the iron-chelating siderophore enterobactin is an example of Iterative NRPSs (Marahiel et al., 1997).

Nonlinear NRPSs (Type C), are characterized by their deviation from the canonically arranged domains $(C\text{-}A\text{-}PCP)_n$. Therefore, at least in one module one or some domains work more than once in the biosynthesis of nonribosomal peptides. Examples for nonlinear NRPSs can be found in the biosynthesis of bleomycin (Mootz et al., 2000) as well as the unusual internal cyclizations or branch-point syntheses in the biosynthesis of vibriobactin or mycobactin (Arnez & Moras, 1997; Conti et al., 1997).

1.4 Hybrid PKS/NRPS

Apart from the NRPS and PKS systems many compounds were shown to be produced by a hybrid of PKS and NRPS assembly lines. The very similar principal for the biosynthesis of secondary metabolites by NRPSs and PKSs, sharing a modular organization in both systems made it possible to couple these two distinct system.

Hybrid NRPS/PKS compounds have been isolated from many of Gram-positive as well as Gram-negative bacteria. Genomic evidence indicates the frequent occurrence of these hybrid systems, which generate many bioactive metabolites (Mizuno et al., 2013;

Fisch, 2013). These secondary metabolites consist often of a polyketide backbone with some incorporated amino acids. The anticancer drug bleomycin is one of the well-known products of such a hybrid system, which consists of 9 NRPS modules and a single PKS module (B Shen et al., 2001).

Two kinds of PKS/NRPS mixed systems were characterized to date. The secondary metabolites of the first class are synthesized from peptide and polyketide moieties, which are assembled separately and then subsequently coupled by a discrete enzyme, e.g. coronatine biosynthesis. In the the second class, interaction between NRPS and PKS modules results in the production of the biosynthetically mixed compound. This class of PKS/NRPS mixed system can be observed in the biosynthesis of myxothiazol, a potent inhibitor of the respiratory chain (Perlova et al., 2006; Rangaswamy et al., 1998).

1.5 Shikimate- derived natural products

The shikimate pathway is an essential biosynthetic route for the biosynthesis of aromatic compounds. This pathway has been found only in microorganisms and plants. In a sequence of seven metabolic steps, phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4P) are converted to chorismate, the precursor of the aromatic amino acids phenylalanine, tyrosine and tryptophan, as well as a number of other aromatic compounds which are critical to sustaining the primary functions of living organisms (Knaggs, 2003; Herrmann & Weaver, 1999).

The final products of the shikimate pathway, as well as all pathway intermediates can be considered branch point compounds that may serve as substrates for other metabolic pathways. Many of the unusual starter units in the process of biosynthesis of natural products are originated from the shikimate pathway.

Some of this shikimate-derived starter units are formed from the end products of the shikimate pathway, i.e. the aromatic amino acids yields phenylpropanoid compounds, e.g. flavonoids or lignans.

A much smaller number of natural products are derived from variants of the intermediates of the shikimate pathway, which branch off at different points (H G Floss, 1997). The following examples show shikimate-derived compounds used as starter units for different secondary metabolites.

Thus cyclohexanecarboxylate (CHC) is the starter unit for the streptomycete antibiotic asukamycin, and several members of the phoslactomycin family (Figure 1-5) (Hu & Floss, 2001; Maeda & Dudareva, 2012).

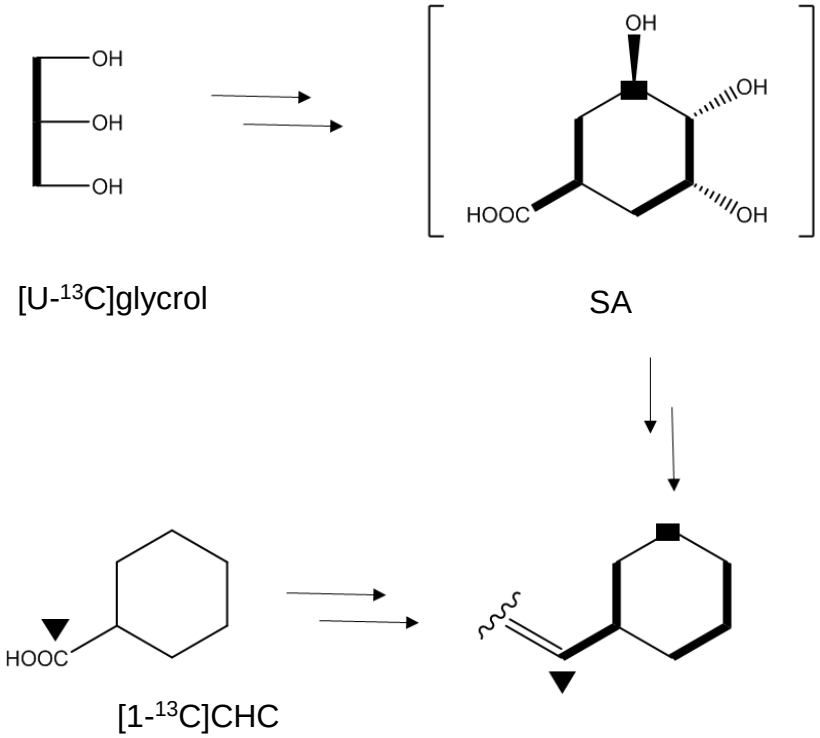


Figure 1-5. The biosynthetic origin of the cyclohexane ring and it`s adjacent carbon shikimic acid (SA)-derived of asukamycin (Moore & Hertweck, 2002).

The antifungal macrolide antibiotic soraphen A from the myxobacterium *Sorangium cellulosum* employs a benzoate-derivative as the starter unit. Feeding experiments verified that the starter unit is derived from glycerol in a manner largely consistent with the shikimate pathway (Hill et al., 2003).

In Enterobacteria three discrete enzymes are involved in *p*-aminobenzoate (PABA) biosynthesis, which may serve as starter unit (Gil & Hopwood, 1983; Gil et al., 1993). PabA and PabB, which share high sequence identity with the *S. griseus bifunctional* PabAB, convert chorismic acid into *p*-amino-*p*-deoxychorismic acid and the lyase PabC converts this intermediate into PABA (Figure 1-6).

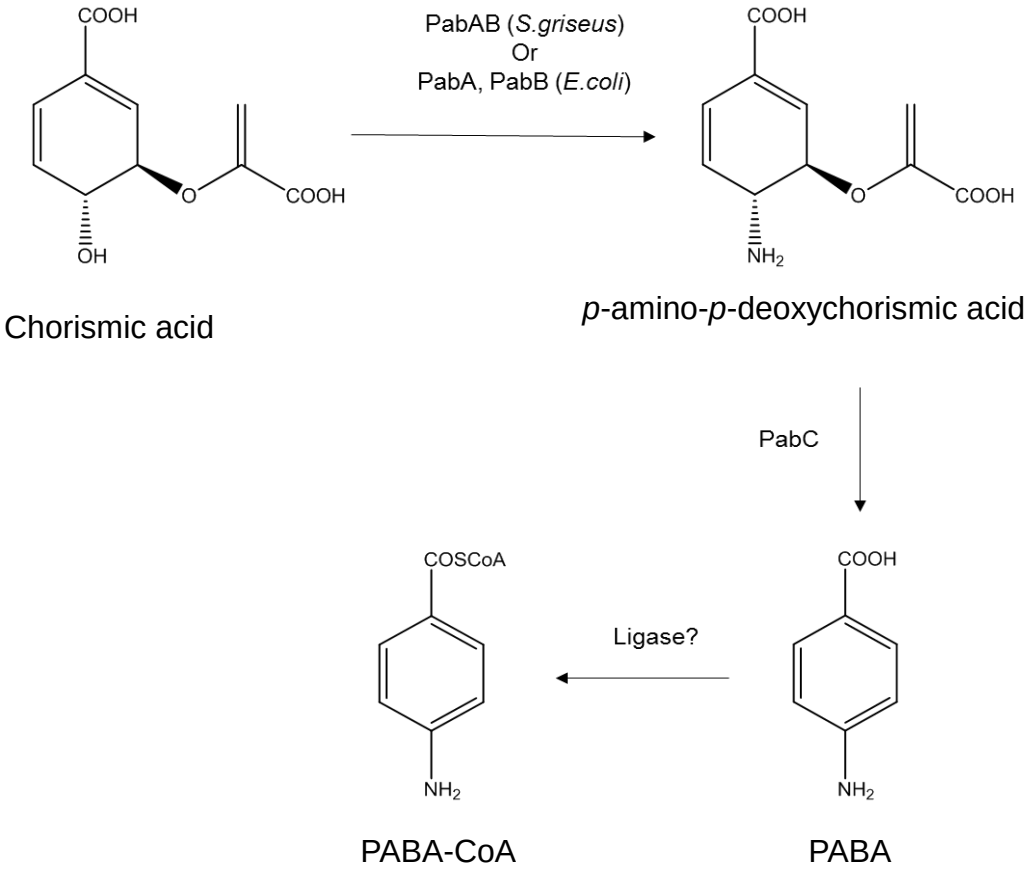


Figure 1-6. The generation of *p*-amino-*p*-deoxychorismic acid from chorismic acid and subsequent conversion to PABA by the lyase PabC (Moore & Hertweck, 2002).

3-Amino-5-hydroxybenzoate (3,5-AHBA), is another type of starter units which originates from the shikimate pathway. The large family of ansamycin antibiotics incorporate a structural element, i.e. mC_7N unit, which is biosynthetically derived from (3,5-AHBA).

The ansamycin class of antibiotics are produced by various actinomycetes. This group of macrolactam natural products are divided into two subgroups according to the structure of the 3,5-AHBA derived aromatic moiety.

The first group, the benzenic ansamycins are mainly cytotoxic against eukaryotes and include ansatrienin A, ansamitocin P3, geldanamycin, and herbimycin A. On the other hand naphthalenic ansamycins, which are active against Gram-positive bacteria as well as *Mycobacterium tuberculosis* include naphthomycin A, tolypomycin Y and rifamycin B.

1.6 Homologues of *dahp* synthase in gene cluster for secondary metabolites

3-deoxy-D-arabinoheptulosonate7-phosphate (DAHP) synthase is the first committed enzyme in the shikimate pathway. This enzyme catalyze the aldol-like condensation of E4P and PEP which results in the generation of 3-deoxy-D-arabino-2-heptulosonic acid 7-phosphate (DAHP) (Furdui et al., 2004). As was described before (See 1-5), the final products and the branched intermediates of the shikimate pathway can be involved in biosynthesis of the natural products. The presence of *dahp* synthase homologues has already been discovered in the biosynthetic gene clusters of secondary metabolites like rifamycin, balhimycin and chloramphenicol. The following context describes the contribution of the DAHP synthase in generation of these secondary metabolites.

Rifamycin (rif) the secondary metabolite from *Amycolatopsis mediterranei*, is the archetype ansamycin, and it is medically important. This strain uses the catalytical activity of DAHP synthase for the generation of the starter unit 3,5-AHBA (Heinz et al., 1999).

The biosynthetic genes for AHBA were first identified as part of the rifamycin biosynthetic gene cluster. Most of the products encoded by these genes are consistent with the proposed AHBA biosynthetic pathway (Figure 1-7), including amino DAHP synthase (RifH), 5-deoxy-5-aminodehydroquinic acid (amino DHQ) synthase (RifG), amino DHQ dehydratase (RifJ) and AHBA synthase (RifK) (Kang et al., 2012).

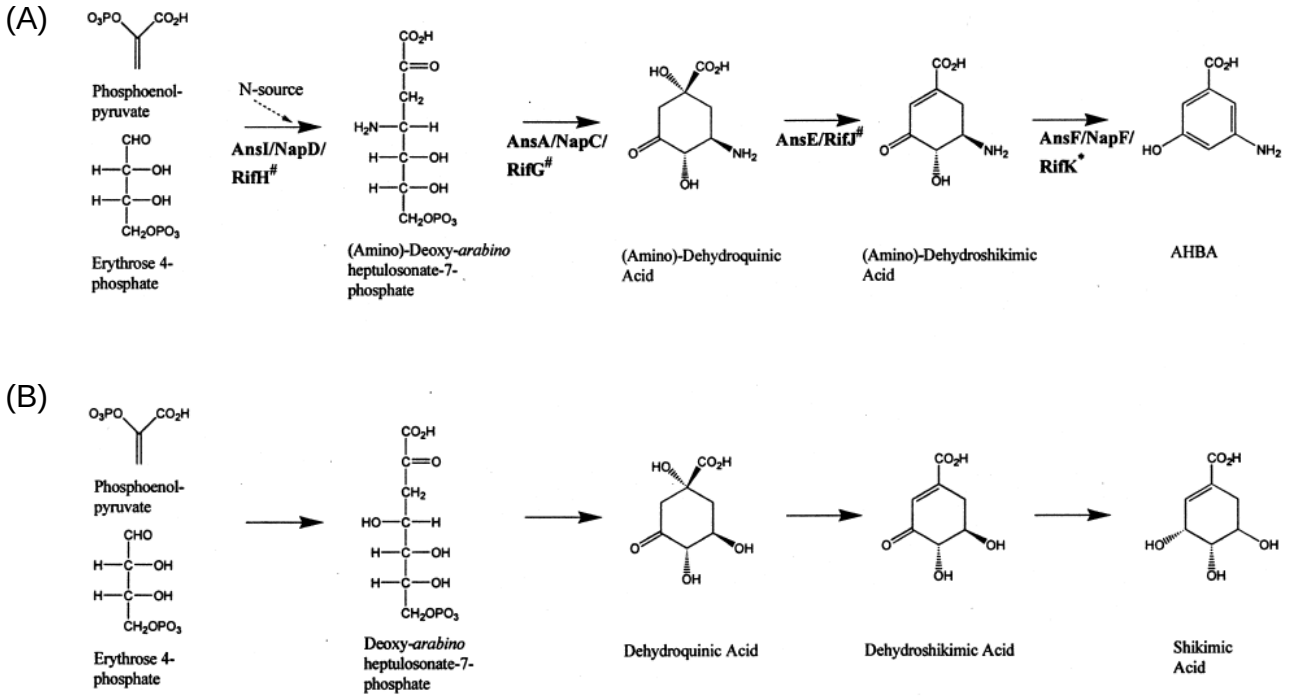


Figure 1-7. Comparison of (A) the proposed AHBA biosynthetic pathway and (B) shikimic acid biosynthetic pathway. Proposed (#) and determined (*) roles of gene products from the rifamycin (*A. mediterranei*), ansatrienin and putative naphthomycin (*S. collinus*) gene clusters are indicated. Required enzymes for the formation of the starter unit AHBA in the biosynthesis of rifamycin: Amino DAHP synthase (RifH), amino DHQ synthase (RifG), amino DHQ dehydratase (RifJ) and AHBA synthase (RifK) (S. Chen et al., 1999).

RifK is involved in the generation of amino DAHP. It was proposed that this enzyme may have two enzymatic activities, a dehydratase activity aromatizing 5-deoxy-5-amino-3-dehydroshikimate, and an aminotransferase activity introducing the nitrogen into a carbohydrate precursor of amino DAHP (T.-W. Yu et al., 2002).

Balhimycin is the secondary metabolite of *Amycolatopsis mediterranei* DSM5908, which belongs to the family Pseudonocardiaceae. The chemical structure of the heptapeptide backbone of balhimycin is similar to the vancomycin (Nadkarni et al., 1994; Pelzer et al., 1999). This glycopeptide antibiotic serve as treatment of severe infections caused by antibiotic-resistant Gram-positive bacteria, in particular by methicillin-resistant staphylococci and penicillin-resistant enterococci (Shawky et al., 2007).

Balhimycin biosynthesis has been elucidated using genetic characterizations of the balhimycin biosynthetic genes, as well as biochemical and chemical analyses by Pelzer and coworkers. The balhimycin biosynthetic gene cluster covers 66 kb and comprises seven complete open reading frames (ORFs) and one incomplete ORF (Pelzer et al., 1999).

Two genes *dahp* and *pdh* which showed high similarity to *dahp* synthase gene from *Amycolatopsis methanolica* and prephenate dehydrogenase from *Zymomonas mobilis* respectively, were identified in the border region of the balhimycin biosynthetic cluster (Figure1-8) (Shawky et al., 2007).

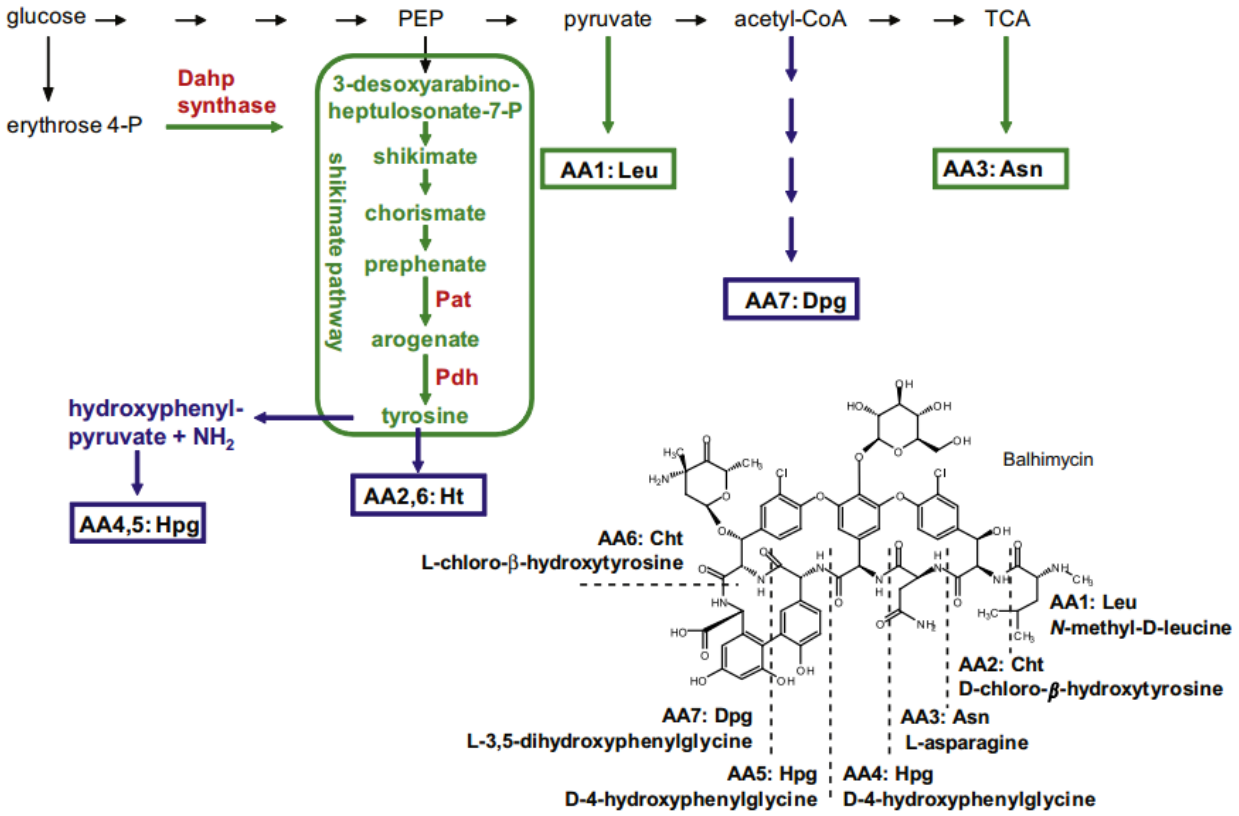


Figure 1-8. Incorporation of the *dahp* and *pdh* genes in balhimycin biosynthesis gene cluster.

Green frames show the necessary reactions or substrates involved in primary metabolism, blue frames Indicate reactions or substrates of the secondary metabolism. Abbreviations: AA, amino acid; Hpg, 4-hydroxyphenylglycine; Dpg, dihydroxyphenylglycine; Ht, hydroxytyrosine; Asn, asparagine; Leu, leucine; PEP, phosphoenol pyruvate; TCA, tricarboxylic acid; Dahp, 3-deoxy-D-arabino-heptulosonate 7-phosphate (Dahp) synthase; Pat, prephenate aminotransferase; Pdh, Prephenate dehydrogenase (Thykaer et al., 2010).

As it was shown in figure 1-8, shikimate pathway either directly or indirectly supplies the b-hydroxytyrosine (b-Ht) at amino acid positions 2 and 6 and hydroxyphenylglycine (Hpg) at amino acid positions 4 and 5 by the aromatic amino acid tyrosine or hydroxyphenylpyruvate as precursor. Therefore, genetic modification of these two genes can be resulted in high or low yield of glycopeptide production (Thykaer et al., 2010).

Thykaer and coworkers performed several experiments by constructing strains expressing an additional copy of the *dahp* gene and a strain carrying an extra copy of both *dahp* and *pdh,* as well as corresponding strains of constructs containing mutants, Δ*dahp* and Δ*pdh*.

The glycopeptide productivities, biomass yields and maximum specific growth rates of these strains were analyzed. The results point out the role of Dahp as limiting factor for high-yield glycopeptide production.

In detail, overexpression of the constructed strains containing either an additional copy of the *dahp* gene or an extra copy of both *dahp* and *pdh* resulted in improved specific glycopeptide production while *pdh* overexpression strain showed a production profile similar to the wild type strain.

On the other hand the production profile of the constructed mutants, Δ*dahp* and Δ*pdh*, were characterized. Deletion of *dahp* resulted in significant reduction in balhimycin production while the production levels of the Δ*pdh* showed a similar pattern as the parent strain (Thykaer et al., 2010).

The *dahp* gene was also identified in other antibiotic biosynthetic clusters, i.e. the chloramphenicol biosynthetic cluster.

The chloramphenicol producing strain *Streptomyces venezuelae,* encodes an enzyme for halogenated natural product bearing an unusual dichloroacetyl moiety that is critical for its antibiotic activity (Schlünzen et al., 2001; Podzelinska et al., 2010).

Chloramphenicol (Figure 1-9), was isolated in 1947 and showed broad range activity against Gram-positive and Gram-negative bacteria, including most anaerobic

organisms. The mechanism of action is based on the inhibitotion of ribosomal peptidyl transferase activity and thereby bacterial growth is inhibited (Izard, 2001; He., 2001).

4-Amino-4-deoxychorismic acid (ADC), an intermediate in biosynthesis of chloramphenicol, is generated from chorismic acid by the encoded protein of *pab*AB.

Sequence analysis of the *S. venezuelae* ISP5230 chromosome by He and coworkers in 2001, has shown that three of four ORFs in the sequence downstream of *pab*AB, an are involved in the shikimate pathway and are necessary for chloramphenicol production (Figure 1-9) (He et al., 2001).

Amino acid sequences of ORF3 and ORF4 were shown to be similar to proteins encoded by monofunctional genes for chorismate mutase and prephenate dehydrogenase, respectively, while the sequence of the ORF5 product resembled DAHP synthase.

The gene *cmlE* which encodes DAHP synthase in the chloramphenicol gene cluster shows the 44% identity to class II DAHP synthases of ansamycin biosynthesis genes in *Streptomyces collinus* Tu1892 as well as DAHP synthases encoding genes within rifamycin biosynthesis in *Amycolatopsis mediterranei* (H. Chen & Walsh, 2001; August et al., 1998).

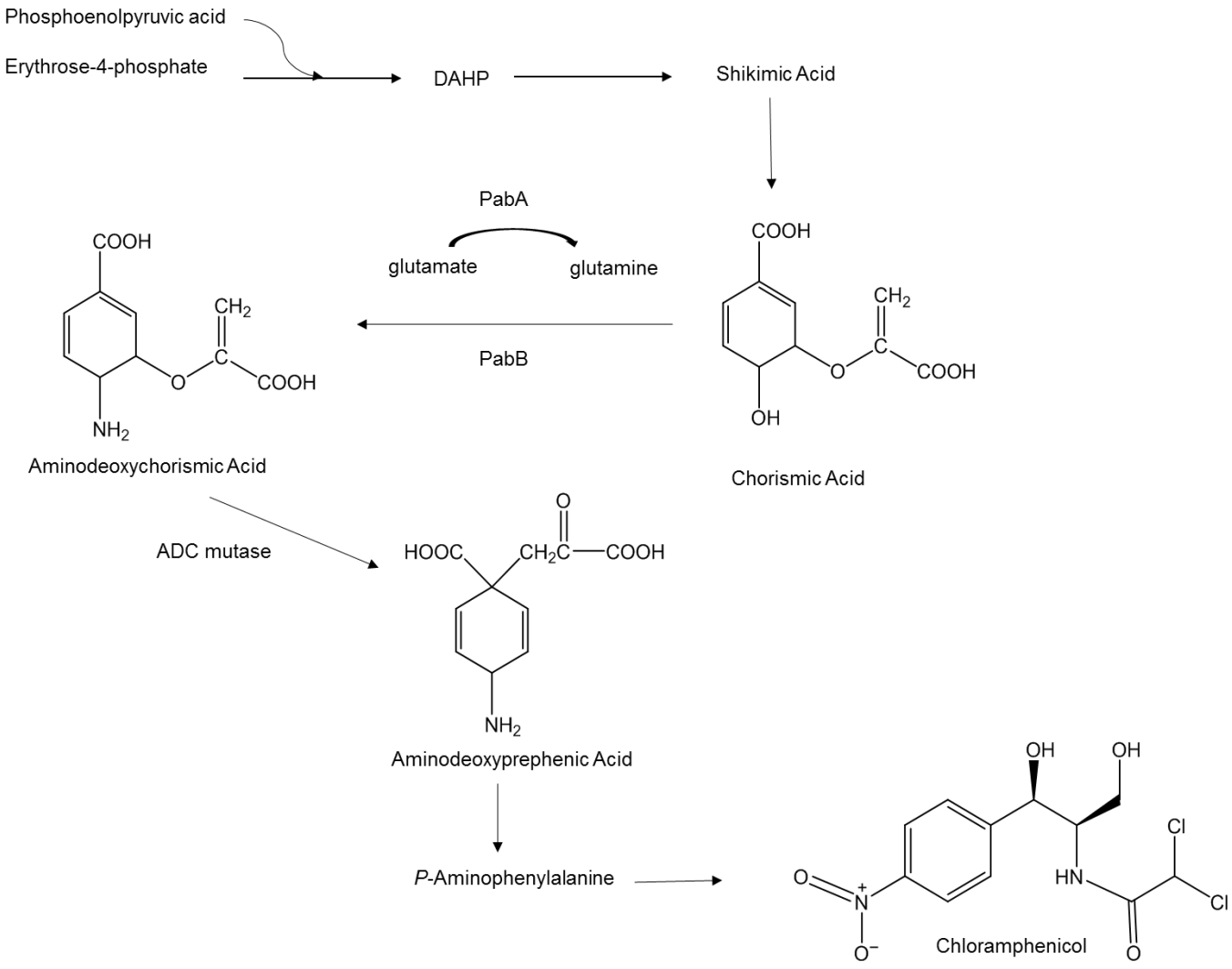


Figure 1-9. Involvement of the shikimate pathway, via chorismic acid in the biosynthesis choloramphenicol in *S. Venezuela* (He et al., 2001).

The following examples show the involvement of DAHP synthase as starter unit provider in the biosynthesis of ambigols A-C (Figure 1-10) and siphonazole (Figure 1-13).

1.7 Ambigols, halogenated secondary metabolites from *Fischerella ambigua*

The ambigol producing strain, *Fischerella ambigua* was originally isolated from a shallow hollow in Mellingen, Kanton Aargau, Switzerland. *F. ambigua* belongs to the phylum Cyanobacteria which are Gram-negative photoautotrophic prokaryotes. Most of the isolated cyanobacterial metabolites are found to be halogenated peptides including anabaenopeptilide 90B, cyanopeptolin 954 (Rouhiainen et al., 2000; Tooming-Klunderud et al., 2007). Ambigols A-C (Figure 1-10) are composed of three simple phenolic building blocks that are linked by either ether-bridges or aryl-aryl bonds to give a trimeric basic structure. These trimeric structure undergoes halogenation which results in incorporation of less than six chloride atoms on the phenyl rings. Ambigols show antialgal, antibacterial, antifungal, molluscicidal, and weak antiviral activities. Antiplasmodial activity was observed for ambigol C while this could not been proven for ambigol A and B (El Omari. PhD thesis).

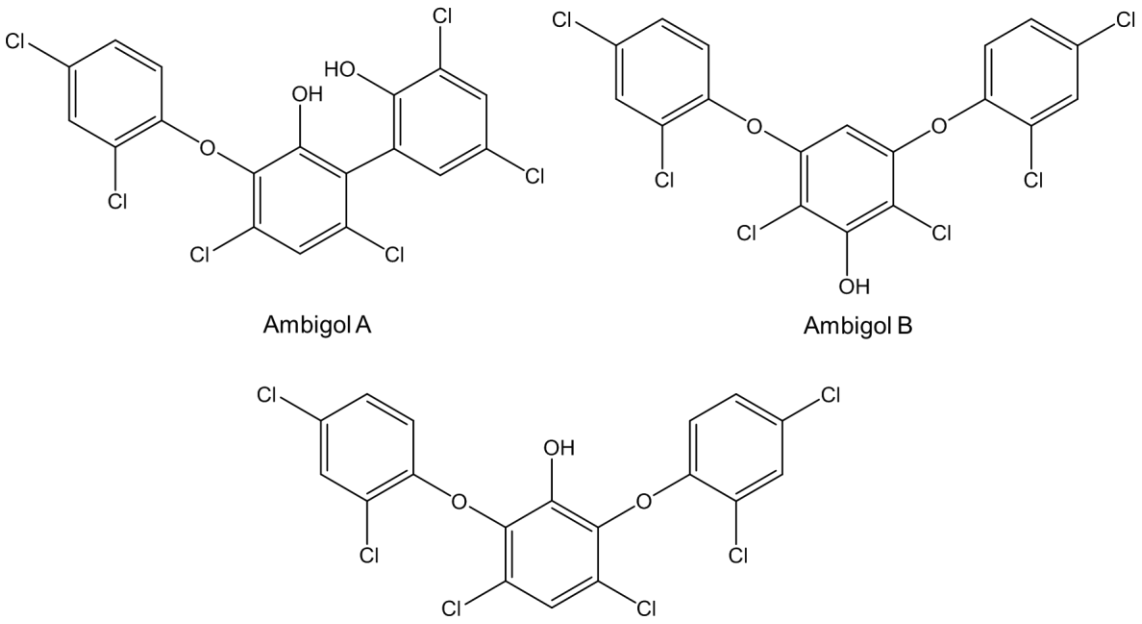


Figure 1-10. Molecular structure of the polychlorinated phenolic ethers Ambigols A-C produced by terrestrial cyanobacterium *F. ambigua*.

1.8 Siphonazole, a hybrid of NRPS/PKS from *Herpetosiphon* sp. 060

Herpetosophon sp. 060 was isolated from a mud sample of the intertidal region (Figure 1-11). *Herpetosiphon* spp. belong to the phylum Chloroflexi which branched off early from the eubacterial stem. The strain presents characteristic physiological features of the Herpetosiphonaceae including gliding motility and filamentous growth (Nett. PhD thesis).

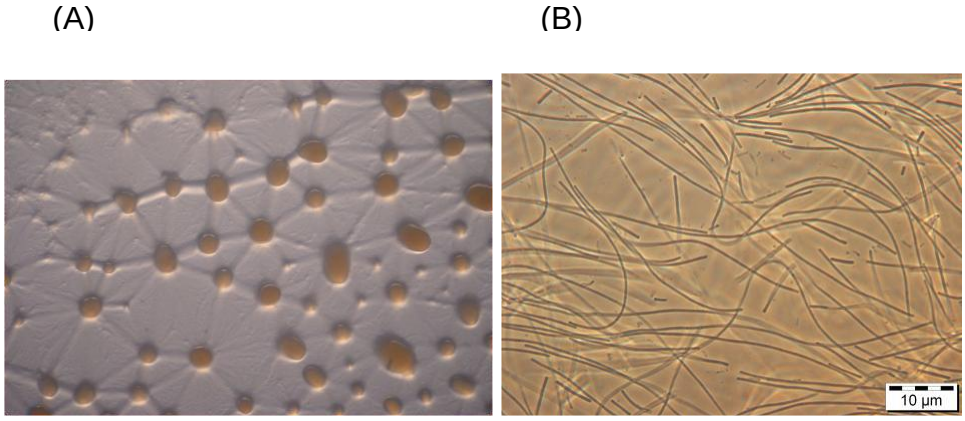


Figure 1-11. Images of *Herpetosiphon* sp. 060. A: orange knob-like colonies on VY/2 agar; B: phase-contrast micrograph (x1000) of *Herpetosiphon* sp. 060 filaments. (Höver. 2012).

Siphonazole, isolated from *Herpetosiphon* sp. 060 (Figure 1-12) shows promising antiplasmodial activity (IC50: 0.59 µg/ml; 1.27 µM). Structure elucidation revealed an unusual composition including a styrene moiety, two oxazole rings connected by a C_2-bridge, and a hitherto unprecedented N-penta[2,4]diene side chain. The origin of all building-blocks of the molecule was elucidated by feeding experiments with labeled precursors which points to: i) hydroxybenzoic acid derived from chorismic acid via the shikimate pathway as precursor, ii) threonine as origin of the oxazole rings as well as a

glycine moiety and acetate units. Labelling studies revealed an unusual dehydration and decarboxylation of the last incorporated acetate unit (Nett. PhD thesis).

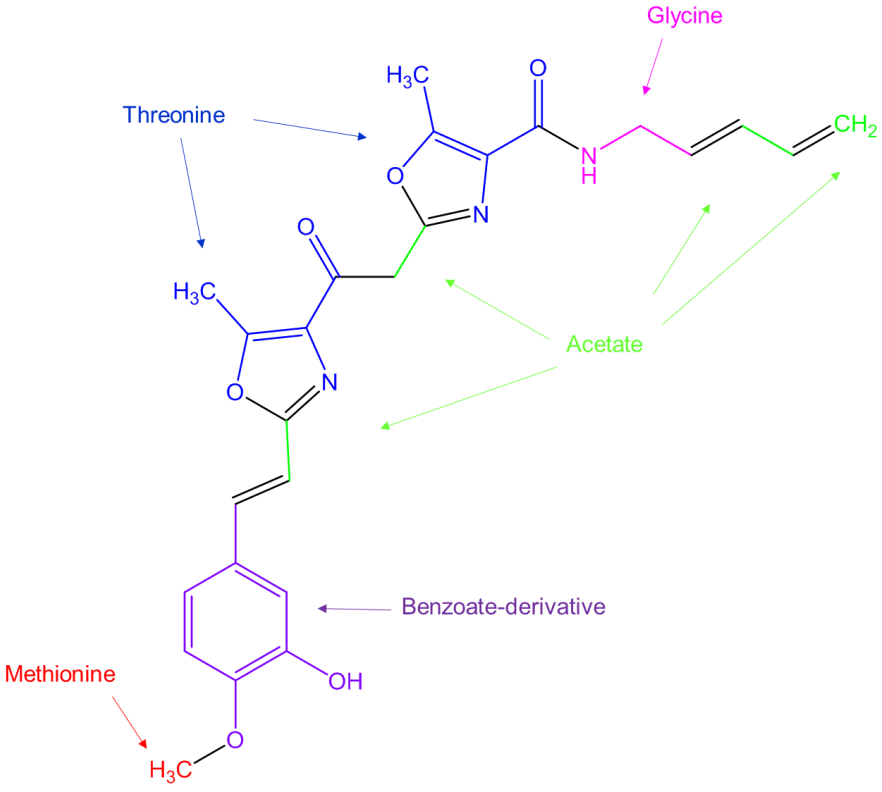


Figure 1-12. Molecular structure of siphonazole and its deduced building blocks.

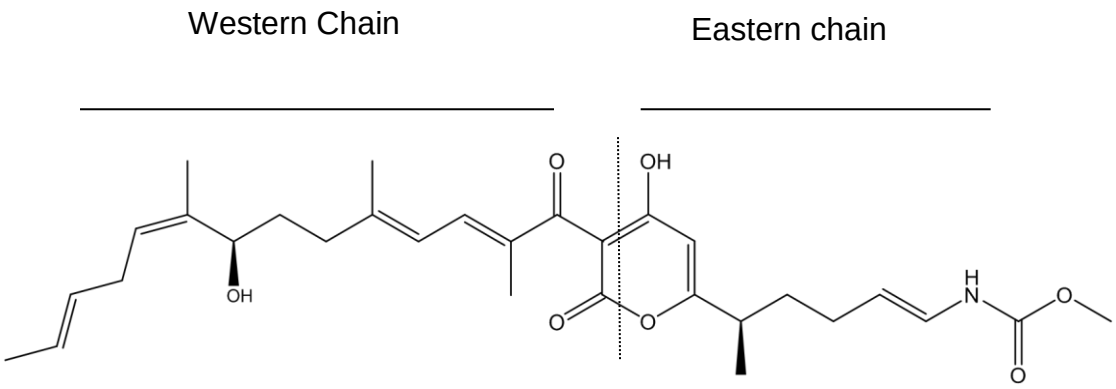
1.9 Corallopyronin A, a myxobacterial compound

Corallopyronin A (Figure 1-13) is produced by strains of *Corallococcus coralloides*, a gliding and fruiting body-forming myxobacterium. Originally identified as an antibacterial substance in the 1980s (Irschik et al., 1985), we recently gained some insights into the biosynthesis of corallopyronin A (Erol et al., 2010).

The mechanistic basis of its antibiotic activity as a noncompetitive inhibitor of the bacterial DNA-dependent RNA polymerase (RNAP) has been investigated (Mukhopadhyay et al., 2008; Belogurov et al., 2009). The molecule inhibits RNAP by binding to a pocket inside of the *holo*-enzyme and thereby interferes with RNA synthesis. RNAP is a proven target for antibacterial treatment, i.e. rifamycins are in clinical use for the treatment of tuberculosis. Corallopyronin A has no cross-resistance toward rifamycins, since it occupies a different binding site (Mukhopadhyay et al., 2008; Belogurov et al., 2009). Corallopyronin A is active against many Gram-positive pathogens, including methicillin-resistant *Staphylococcus aureus*, but is most promising against infections in which intracellular bacteria are involved, i.e. *in vivo* activity against *Wolbachia*, the obligate endosymbionts of filarial nematodes was proven by us (Schiefer et al., 2012). Targeting these bacteria was shown to be a good approach to control filarial infections, which cause elephantiasis and river blindness (Bockarie et al., 2009; Hoerauf et al., 1999), the latter being leading causes of morbidity in developing countries. For the development of further antibiotics detailed investigation of the interactions between the RNAP and its different inhibitors is important, especially when considering the quick development of resistance against RNAP inhibitors (Campbell et al., 2001).

Corallopyronin A is synthesized by a multi-enzyme complex that is a combination of a PKS and NRPS. In general, this type of enzyme is well characterized and acts in a modular way, combining key biochemical steps in an assembly line fashion to create structural diversity (Hertweck, 2009). The biosynthesis of corallopyronin A, however, is

atypical and has several peculiarities. The molecule is synthesized from two separate chains (eastern and western chains), which are finally intermolecularly connected to form the precursor, i.e. precorallopyronin A, of the active compound.



Corallopyronin A

Figure 1-13. Structure of corallopyronin A. Parts of precorallopyronin, originating from the western and eastern chains, respectively, are separated by the dotted line.

2 Scope of the Study

In this study further insights into the biosynthesis of bioactive secondary metabolites from gliding bacteria should be obtained. In particular, parts of the biosynthetic gene clusters belonging to, (i) ambigols from the cyanobacterium *Fischerella ambigua*, (ii) siphonazole from *Herpetosiphon* sp. 060, belonging to the Chloroflexi, and (iii) corallopyronin A produced by the myxobacterium *Corallococcus coralloides* B035, were envisaged to be studied.

Using bioinformatic tools, the biosynthetic gene cluster of the ambigols, siphonazole and corallopyronin A had been elucidated, and showed that biosyntheses of the mentioned natural products have several peculiarities. Therefore, the aim of the present study was to obtain a deeper understanding of the biosynthesis of these secondary metabolites, i.e. the formation of starter units in the biosynthetic processes.

Presence of a 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase sequence in the ambigols, as well as in the siphonazole biosynthetic gene clusters was suggested in previous studies (Höver. PhD thesis & El. Omari PhD thesis). Therefore, the analysis of the function of these putative enzymes was a goal of the current study.

Concerning corallopyronin A, the function of the CorI loading module and of the *O*-methyltransferase CorH, the latter hypothetically involved in formation of the unusual starter unit carbonic acid methyl ester, should be probed.

In another part of the study the biosynthesis of siphonazole by heterologous expression of the gene cluster in Bacilli or Streptomycetes should be addressed. Once a recombinant strain would be available, genetic manipulation was planned to shed light on several peculiarities in siphonazole biosynthesis, like the unusual decarboxylation and dehydration of the last incorporated acetate unit.

3 Material and Methods

3.1 Solvents and Reagents

Solvents and reagents were obtained from Roth (Karlsruhe, Germany), Sigma-Aldrich (Steinheim, Germany) or Fluka (Taufkirchen, Germany).

3.2 Enzymes

The enzymes used in this work were purchased from Fermentas (St. Leon Rot, Germany), Promega (Mannheim, Germany) or Roth (Karlsruhe, Germany). They were applied according to the respective manufacturers´ instructions for use.

3.3 Media, stock solutions and buffers

Media	Ingredients
LB medium	10 g tryptone, 5 g yeast extract, 10 g NaCl, water ad 1000 mL, pH 7.5
LB agar	10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar, water ad 1000 mL, pH 7.5
MD1 + G medium	3 g casitone, 0.7 g $CaCl_2$ x $2H_2O$, 2 g $MgSO_4$ x $7H_2O$, 2.2 g glucose, water ad 1000 mL, pH 7.5; after sterilization 1 mL trace element solution I and 1 mL vitamin B12 solution were added
SOB medium	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 0.186 KCl, water ad 1000 mL, pH 7.5

SOC medium	20 g tryptone, 5 g yeast extract, 0.5 g NaCl, 0.186 KCl, water ad 1000 mL, pH 7.5, after sterilization 1 mL of 1 M glucose solution was added
NCM	17.4 g K_2HPO_4, 11.6 g NaCl, 5 g Glucose, 5 g Tryptone, 1 g Yeast extract, 0.3 g Trisodium citrate, 0.05 g $MgSO_4.7H_2O$, 91.1 g Sorbitol, water ad 1000 ml, pH 7.2 (G. Q. Zhang et al., 2011)

Stock solutions	**Ingredients**
Trace element solution I	20 mg $ZnCl_3$, 100 mg $MnCl_2$ x $4H_2O$, 10 mg H_3BO_3, 10 mg $CuSO_4$, 20 mg $CoCl_2$, 5 mg $SnCl_2$ x $2H_2O$, 5 mg LiCl, 20 mg KBr, 20 mg KI, 10 mg Na_2MoO_4 x $2H_2O$, 5.2 g Na-EDTA x $2H_2O$, water ad 1000 mL
Vitamin B12 solution	0.5 mg cyanobobalamine ad 1 mL water
Ampicillin stock	100 mg ampicillin ad 1 mL water
Apramycin stock	100 mg apramycin ad 1 mL water
Chloramphenicol stock	12 mg ampicillin ad 1 mL ethanol
Kanamycin stock	60 mg kanmycin ad 1 mL water
Streptomycin stock	100 mg streptomycin ad 1 mL water

Zeocin stock	25 mg streptomycin ad 1 mL water

Buffers	**Composition**
Buffer P1	50 mM Tris- HCl (pH 8), 10 mM EDTA, 100 µg/ml RNase A
Buffer P2	200 mM NaOH, 1 % SDS
Buffer 3	3 M potassium acetate (pH 5.5)
TE-buffer	10 mM Tris- HCl (pH 8), 1 mM EDTA
10x TBE-buffer	0.89 M Tris, 0.02 M EDTA, 0.87 M H_3BO_3
Protein lysis buffer	50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8
Protein wash buffer	50 mM NaH_2PO_4, 300 mM NaCl, 30 mM imidazole (alternative 50 mM), pH 8
Protein elution buffer	50 mM NaH_2PO_4, 300 mM NaCl, pH 8; for gradual elution use 100, 150, 200, 300 mM imidazole
10x glycine SDS electrophoresis buffer	250 mM Tris, 2 M glycine, 1% SDS, pH 9

Staining solution	10% acetic acid, 50% ethanol, 0.005% coomassie brilliant blue R-250, 40% water
Destaining buffer	10 % acetic acid, 30 % methanol, 60% water
ETM buffer	0.5M Sorbitol, 0.5 M Mannitol, 10% Glycerol

3.4 Kits and standards

Gene Ruler™ DNA Ladder Mix	Fermentas GmbH (St. Leon-Rot, Germany)
Gene Ruler™ 1kb plus DNA ladder	Fermentas GmbH (St. Leon-Rot, Germany)
PageRuler™ Unstained Broad Range Protein Ladder	Thermo Scientific™(Schwerte, Germany)
PageRuler™ Unstained Low Range Protein Ladder	Thermo Scientific™(Schwerte, Germany)

QIAquick PCR Purification Kit	Qiagen GmbH (Hilden, Germany)
QIAquick Gel Extraction Kit	Qiagen GmbH (Hilden, Germany)
Wizard® SV Gel and PCR Clean-Up System	Promega (Mannheim, Germany)
Wizard® Genomic DNA Purification Kit	Promega (Mannheim, Germany)
PureYield™ Plasmid Miniprep System	Promega (Mannheim, Germany)
DNA Clean & ConcentratorTM	Zymo Research Corporation (Irvine, USA)
Zymoclean™ Large Fragment DNA Recovery Kit	Zymo Research Corporation (Irvine, USA)

3.5 Organisms

Organism	Genotype	Provider
XL1-Blue *E. coli* cells	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10* (Tetr)	Stratagene (La Jolla, CA,USA)
BL21 *E. coli* cells	*F- ompT gal dcm Ion hsdSb (rB- mB-) λ(DE3 [lac lacUV5-T7 gene1 ind 1 sam7 nin5])*	Invitrogen Life Technologies Corporation (Karlsruhe, Germany)
BAP1 *E. coli* cells	n.a.; *sfp* from *B. subtilis*	Pfeifer et al., 2001
ET12567 *E. coli* cells	*Dam-, dcm-, hsdM-*	MacNeil et al., 1992
BW25113 *E. coli* cells	*F-, Δ(araD-araB)567, ΔlacZ4787 (::rrnB-3), λ-, rph-1, Δ(rhaD-rhaB)*568, *hsdR*514	(Gust et al., 2003)
TOP10 *E. coli* cells	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG)*	Invitrogen Life Technologies Corporation (Karlsruhe,Germany)
Bacillus amyloliquefaciens FZb42	Wild Type	Prof. Piel lab

3.6 Vectors

Vector	Resistance	Manufacturer
pET28a(+)	Kanamycin	Merck KGaA (Darmstadt, Germany)
pCC1FOS™	Chloramphenicole	EpicentreBiotechnologies (Madison, U.S.A)
pGEM-T	Ampicillin	Promega (Mannheim, Germany)
pIJ773	Apramycin	(Gust et al., 2003)
pIJ778	Streptomycin	(Gust et al., 2003)
pIJ790	Chloramphenicole	(Gust et al., 2003)
pKD46	Ampicillin	(Datsenko and Wanner., 2000)
pGF27	Zeocin	Dr. Hartzell lab
pESAC 13	Kanamycin	Bio S&T
pET151/D-TOPO®	Ampicillin	Invitrogen Life Technologies Corporation (Karlsruhe, Germany)

3.7 DNA-Constructs

The following table lists the DNA-Constructs used or made In this study.

Construct	Vector	Inserte
EC9	pCC1FOS	Genomic DNA from *Herpetosiphon* sp.060 with parts of the putative siphonazole gene cluster
EC10	pCC1FOS	Genomic DNA from *Herpetosiphon* sp. 060 with parts of the putative siphonazole gene cluster
EC10+*bae*j	pCC1FOS	Genomic DNA from *Herpetosiphon* sp. 060 with parts of the putative siphonazole gene cluster +*baeJ*
IC2	pCC1FOS	Genomic DNA from *Herpetosiphon* sp. with parts of the putative siphonazole 060 gene cluster
pCC1FOS-siphonazole	pCC1FOS	Fosmid with complete putative siphonazole cluster, recombined from EC9+IC2 and EC10
pCC1FOS-siphonazole +*baej*	pCC1FOS	Fosmid with complete putative siphonazole cluster and *baej* gene

pET151+ACP1	pET151	*corI* ACP
pET28CorH*	pET28	*corH o*-MT
pETDAHP Fa	pET28 (+)	*amb7* DAHP synthase
pET151- SphI	pET151	*sphI* DAHP synthase
pET151_SphJ	pET151	*sphJ* TE
pGF27+*baeJ*	pGF27	*baeJ*
SphH Hyd-ACP	pET151	*sphH* Hyd+ ACP

3.8 Primers

Primers used in this work were designed using CloneManager 9 and purchased from Eurofins MWG Operin (Ebersberg, Germany). The oligonucleotides were reconstituted in sterile water and adjusted to a concentration of 100 pmol/µl. These primers were stored short-term and long-term at 4°C and -20°C respectively. A list of the sequences of primers used in this work is given in the appendix (7.1).

3.9 Software and databases

Basic Local Alignment Search Tool (BLAST) [http://www.ncbi.nlm.nih.gov/pubmed] is provided by the National Center for Biotechnology Information (NCBI). This software was used for multiple sequence alignment for protein primary sequences and for

nucleotide sequences. Blastp was used for protein homologies and for analyzing conserved domains. Blastn was used to compare a nucleotide sequence with the nucleotide database of NCBI.

Chem Draw Pro 14.0 this program was used for sketching chemical structures as well as determination of their exact mass and molecular weight.

Chromas Lite 2.01 this software was used for analyzing and editing the DNA sequences.

CloneManager 9 this system helps with primer designing and for planning of cloning experiments, open reading frame analysis, as well as preparing high-resolution graphic maps. **Restriction cleavage tool** [http://tools.neb.com/NEBcutter2/] was used for carry out restriction side analysis.

ClustalW [http://www.ebi.ac.uk/Tools/clustalw] is provided by the European Bioinformatics Institute (EBI), (EMBL). This tool was used to generate multiple sequence alignments of amino acid sequences.

DNAtrans [http://www.b-und-s-software.de/] is an easy to use tool for cleanup of sequences, primer check, and for translation.

Kodak 1D software version 3.5.4 and iX Imager software these programs were provided together with the respective gel documentation system (Kodak Scientific Imaging Systems and Intas Gel iX imager) and were applied to edit digital photos of electrophoresis gels.

Graf Pad Prism 5 this program was used to determine kinetic constants and to create the graphs.

Pfam provided by Sanger institute is a database of protein families that includes their annotations and multiple sequence alignments generated using hidden Markov models.

3.10 General molecular biological methods

3.10.1 Sterilization

Mediums, solution and all the microbiological working tools were sterilized by autoclaving for 20 min at 121°C and 2 bars pressure in a Varioklav steam sterilizer. Heat sensitive solutions like antibiotics were sterilized by filtration through the 0.22 µm membrane.

3.10.2 Storage and disposal of organisms

For the long term storage of the microorganisms, the glycerol cultures were prepared in cryogenic vials. 750 µl of the overgrown bacterial culture were mixed by the same amount of the sterile 100% glycerol and stored at -80°C. For the short-term storage the same vials were stored at -20°C and used for inoculation of the mediums.

For waste disposal of microorganisms they were autoclaved for 20 min at 121°C and 2 bar.

3.10.3 Cultivation of organisms

All the working steps regarding the cultivation of the microorganisms were performed under the sterile condition on a laminar air flow clean bench. For the small volume of *E.coli* cultures the 10 ml tubes or 2 ml Eppendorf tubes were used. For big scale cultures 100-5000 ml Erlenmeyer flasks were used. For the cultivation on agar plates 100-200 µl of the bacterial suspension were poured and spread on the plate and dried under the sterile bench. The liquid cultures of *E. coli* were incubated on a horizontal shaker at 16-37°C and 140 rpm.

3.11 Molecular biological methods concerning bacterial organisms

3.11.1 Preparation of chemically competent *E. coli* cells

A single *E.coli* colony was inoculated in 3 ml LB medium and incubated overnight at 37°C and 160 rpm. The grown cells were inoculated in the 300 ml flask containing 70 mL of 2xYT medium. The bacterial suspension was incubated at 37°C and 180 rpm until an OD_{600} of approximately 0.3-0.4 was reached. Later, cells were harvested by centrifugation at 8000 ×g for 10 min at 4°C. Supernatant was decanted and the pellet resuspended in about 10 mL of ice cold $CaCl_2$/$MgCl_2$-solution (70 mM $CaCl_2$, 20 mM $MgCl_2$) and kept cold in ice for 30 min. Cells were harvested anew and suspended in 3.5 mL ice cold $CaCl_2$/$MgCl_2$ solution and 875 µL glycerol. The competent cells were aliquoted into sterile 1.5 mL microfuge tubes and stored in the -80°C freezer.

3.11.2 Preparation of electro-competent *E. coli* cells

Electro-competent cells were always freshly prepared and used at the same day. All the steps were carried on at 4°C.

100 ml of the SOB medium was inoculated with 3 ml of the *E.coli* cells which were grown over night at 30°C. The new inoculated medium was incubated at 30°C until OD_{600} was equal to 0.5. The culture was harvested by centrifugation, at 6000 rpm for 5 min. The supernatant was then discarded and the pellet was treated in 25 mL 10% glycerol solution. The washing step was performed once with 20 mL, then with 10 mL and in the end with 5 mL of the same solution. Finally the pellet was resuspended in 300 µl of 10% glycerol solution which then stored in 100 µl microfuge tubes.

3.11.3 Preparation of electro-competent *Bacillus* cells

The procedure for preparing the *Bacillus* electrocompetent cells was performed after (G. Q. Zhang et al., 2011).

A single colony of *Bacillus amyloliquefaciens* FZb42 was inoculated in 3 ml NCM medium and shaken overnight at 37°C and 200 rpm. The pre-culture was used for inoculation of 300 ml NCM medium in 1/100 ratio. The main culture was incubated at 37°C and monitored by measuring the optical density until an OD_{600} reached 0.5. In order to obtain cell-wall weakening and/or cell-membrane fluidity disturbance, a solution of Gly, $_{DL}$-Thr and Tween 80 (3.89% of Gly, 1.06% of $_{DL}$-Thr, and 0.03% of Tween 80) was added to the to the culture and then the incubation time was extended for an extra one hour. The cell culture was cooled on ice for 20 min, and harvested by centrifugation at 4°C, 6000 rpm for 5 min. The pellet were washed four times with ice-cold ETM buffer (0.5 M sorbitol, 0.5 M mannitol, and 10% glycerol), and the electro-competent cells were resuspended in 1/100 volume of the original culture. After the final washing step, the pellet was resuspended in 1 ml of the ETM buffer and kept in ice until to be used in the experiment.

3.11.4 Heat shock transformation of *E.coli* cells

In order to *in vivo* amplification of plasmids, protein expression or λ-Red recombination *E. coli* cells were transformed with foreign DNA. The 100 µL aliquot of chemical

competed cells were thawed on ice and mixed with 5-10 µl of DNA. The mixture was incubated in ice for 30 min, then switched to 42°C for 90 s and transferred back again to the ice. The cells were recovered by adding 1 ml of LB medium incubated for 1 hour at 37°C to allow expression of the antibiotic resistant genes. Afterward, 100 µl of the suspension was spread on the LB agar supplemented with the suitable antibiotic, while the rest was harvested by 1 min centrifugation at 8,000 ×g. Most of the supernatant containing the medium was discarded and the cell pellet was resuspended in the last remained drops and spread on the agar plates. These plates were incubated overnight at 37°C. In the case of the pGEM-T vector, the agar plates were additionally supplimented with x-gal (4 mg/mL) for selection via blue-white screening.

3.11.5 Transformation of *E.coli* cells by electroporation

For high transformation efficiency and also in the case of transferring the large fragments, electroporation was used in this study. For this purpose, the freshly prepared electro-competent cells were used. Depending on the experiment, the cells were mixed with 5-15 µl of the DNA to be introduced and loaded into the pre-chilled electroporation cuvette with a diameter of 1 mm. Afterward, a voltage of 25 kV / 1cm was applied in the Biorad MicroPulser™. Cells were immediately recovered by 1 ml SOB medium and incubated in 2 ml microfuge tubes at 30°C for 1h. Cells were harvested by 1 min centrifugation at 6,000 ×g, most of the supernatant discarded and pellet was resuspended in the last drops. The suspension, later, was spread on the agar plates containing the appropriate antibiotic for selection and incubated overnight at 30°C.

3.11.6 Transformation of *Bacillus* cells by electroporation

Electro-competent cells (100 µl) were mixed on the ice with 3-7 µl of the plasmid DNA and filled into ice-cold 1-mm gap electroporation cuvettes. After a brief incubation on ice, the mixture was shocked by a single 2.1 kV/cm pulse generated by Biorad MicroPulser™. The cells were immediately diluted into 1 ml of the corresponding

recovery medium (NCM medium plus 0.38 M mannitol) and incubated at 37°C for 3-7 hours. Afterward, the cell suspension was spread on the LB agar plates containing the appropriate antibiotic for the selection.

3.11.7 Preparation of cell lysis from *Corallococcus coralloides*

To perform the ACP1 loading assay, the cell lysis from *C. coralloides* was required. Therefore a 100 ml culture of *C. coralloides* B035 was grown in MD1+G medium for 2 days at 30°C. The cells were harvested by centrifugation and the pellet resuspended in 5 ml of lysis buffer. In order to breaking the cells, the bacterial suspension was subjected to ultrasonic treatment. The yielded crude extract was centrifuged for 30-45 min at 10,000 ×g and 4°C and. The supernatant was always prepared fresh and directly applied as cell lysate in the assay.

3.12 Molecular biological methods concerning nucleic acids

3.12.1 Isolation of genomic DNA

To amplify the targeted genes, genomic DNA of *C.coralloides* B035, as well as *Bacillus amyloliquefaciens* FZb42 was needed. The genomic DNA of these strains was isolated using the Wizard® Genomic DNA Purification Kit, according to the manufacturers' instructions.

The isolated genomic DNA used in this study for amplification of *amb7* was provided by Dr. Mustafa El Omari.

3.12.2 Isolation of plasmid/fosmid DNA

To isolate plasmid or fosmid, a single colony of the *E.coli* cells harboring the corresponding construct was inoculated in 3 ml medium and then incubated overnight. Subsequently, the vectors were purified with the Promega´s PureYield Miniprep, according to the manufacturers' instructions.

For larger demands of the DNA vector, 50-100 ml of medium was inoculated and the vectors were isolated using the same amount of P1, P2 and P3 buffers. In the first step, the bacterial suspension was harvested by centrifugation and suspended in buffer P1 (6.5 ml for 50 ml culture). To open the cells, P2 buffer was added to the mixture and incubated at room temperature for 5 min. Precipitation was enhanced by using ice-cold neutralized buffer (P3) and centrifugation. The supernatant containing the DNA was purified by adding equal amount of phenol and chloroform. The aqueous phase was recovered and the DNA was precipitated with isopropanol and centrifugation. The obtained pellet was washed by 70% ethanol, dried at 60°C and, finally, solved in 500 µl Nuclease-Free Water.

3.12.3 Polymerase chain reaction

The polymerase chain reaction (PCR) is certainly the most useful laboratory technique in molecular biology. PCR allows for exponential amplification of sequence-specific targets between two primer regions in a DNA molecule. The PCR consists of a series of 20 to 40 repeated temperature changes, called cycles. Cycling is often preceded by a single temperature step to denaturize the DNA by heating so primers can anneal to their complement sequence of the template. Using the DNA-polymerase, the elongation of the new DNA strands is performed. Repetition of the mentioned cycle results in amplification of the several series of the targeted fragment.

Most commonly, *Thermus aquaticus* (Taq) derived DNA-polymerase is used for PCR reaction. The amplification of an open reading frame used for heterologous protein

expression was performed with DNA-polymerase from *Pyrococcus furiosus* (Pfu), which is capable of proofreading activity.

A standard PCR reaction mixture was composed as follows:

10x PCR buffer	4 µL
10x MgCl2-solution (25 mM)	1 µL
DMSO	1 µL
Primer 1 (100 µM)	0.5 µL
Primer 2 (100 µM)	0.5 µL
dNTPs (10 mM)	0.4 µL
Taq polymerase (5 U/µL)	0.16 µL
DNA template	1 µL
Water	ad 20 µL

Thermocycling conditions for a routine PCR is described below:

1. Initial denaturation	95°C	5 min	
2. Denaturation	95°C	30 s	30×
3. Annealing	53-65°C	30 s	
4. Elongation	72°C	30-120 s	
5. Final elongation	72°C	4-5 min	
6. Cooling	4°C	hold	

Annealing temperature was calculated to be about 5°C below the melting temperature of the primers. The elongation periods were chosen according to the size of expected product. Steps 2-4 were repeated 30 times before proceeding with step 5. To avoid the unspecific amplification, a negative control without the template was used.

3.12.4 Whole cell PCR

For rapid screening of the bacteria which transferred with the foreign DNA, whole cell PCR was used. Cells on the agar plates were picked with sterile tooth picks and stirred in to the PCR mixture. In the case of bacterial suspensions, 1-2 µl of the culture was applied as template. The initial denaturation (3.12.3) was prolonged to 10 min, in order to certify the denaturation of bacterial cell wall and DNA.

3.12.5 Sequencing of DNA constructs and PCR fragments

To avoid mutation during the amplification of the targeted genes, the nucleic acid sequence of the PCR product was analyzed. Sequencing was performed by GATC Biotech AG (Konstanz, Germany) after the sanger dideoxy or chain termination method (Sanger et al., 1977). Therefore the yielded PCR product was cloned in the pGEM-T vector which served as the DNA template for sequencing. This template was denatured and amplified using one primer by PCR in four separate sequencing reactions. Each contained dNTPs, as well as one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP) which integrated in the growing DNA chain. Obtained DNA fragments were denaturized and separated by size using gel electrophoresis. The DNA bands were visualized by autoradiography or UV light and the DNA sequence were elucidated using X-ray film or gel image.

3.12.6 Restriction digestion

PCR products or plasmid DNA samples were subjected to restriction digestions with restriction endonucleases. These enzymes belong to the bacteria and archaea defense mechanism against invading viruses. They catalyze the cleavage of the sugar-phosphate backbone of both DNA strands, within the specific palindromic sequences with 4–8 bp length. Due to the palindromic nature of restriction sites, these fragments can be ligated to DNA cut with either the same enzyme or another that produces compatible ends. To obtain the clear orientation of incorporated DNA fragments in vectors, the DNA was restricted by two different restrictions. In the case of using one restriction enzyme, the linearized DNA was subjected to dephosphorylation, in order to avoid the self-ligation.

Restriction enzymes in this work were applied according to the manufacturers' instructions.

3.12.7 Dephosphorylation of linear DNA

Vectors restricted by a single restriction enzyme, are more likely to re-ligate back on itself than on an added DNA fragment. Therefore to decrease the background activity of the cloning process through the self-ligation, vectors were subjected to dephosphorilation. Dephosphorylation was accomplished by shrimp alkaline phosphatase (SAP) or calf intestinal alkaline phosphatase (CIAP) according to the manufacturers' instructions. These enzymes catalyze the removal of 5' phosphate of the linearized vector prior to ligation. The enzyme was heat-inactivated and removed from the DNA mixture by Zymo Research DNA Clean & Concentrator™ kit.

3.12.8 Ligation of DNA fragments

Ligation is an essential laboratory procedure in molecular biology whereby recombinant DNA molecules are generating from joining DNA fragments. This was used in this study

to introduce the foreign DNA fragments into suitable vectors. Thus, the obtained linear DNA fragments of PCR products or vectors by restriction digestion (3.12.6) were subjected to ligation using T4 DNA-ligase. This enzyme catalyze the linkage of the end parts of DNA fragments, by forming of phosphodiester bonds between the 3' hydroxyl and 5' phosphate of adjacent DNA residues. To test the accuracy of the resulted construct, the ligation mixture was transferred to the *E.coli* XL Blue cells either by heat-shock transformation (3.11.4) or by electroporation (3.11.5). The isolated plasmids from the corresponding *E.coli* Cells were used to prove the integration of the insert into the vector by PCR and by restriction digestion.

The standard ligation mixture contained:

10x T4 DNA Ligase Buffer	2 µl
T4 DNA ligase	1 µl
Vector DNA	5 µl
Insert DNA	8 µl
Nuclease-free Water	ad 20 µl

The reaction mixture was incubated overnight at 16°C, and afterward the enzyme was heat- inactivated at 65°C for 10 minutes. In the case of ligation of the big fragments, PEG- mediated ligation was used. Up to 5% (w/v) high-quality PEG 8000 was added to the reaction mixture and reaction mixture was incubated at 4°C overnight. PEG 8000 stimulates ligation by increasing macromolecular crowding and aggregation of DNA molecules. Under this condition, ligation efficiency can be improved by approximately three- to six-fold (Sambrook & W Russell, 2001).

3.12.9 Agarose gel electrophoresis

Gel electrophoresis is a tool for analyzing the DNA fragments, which can be used for separating DNA fragments by size as well as for visualization and purification.

Therefore, in this study, the restricted DNA fragments, as well as PCR products, were evaluated according to their size by agarose gel electrophoresis and when it was needed, the desired band was extracted from the gel. Depending on the size of fragment, gels with a concentration of 0.7% – 3% in 1x TBE-buffer were used. Except for the PCR products which were directly loaded on the gel, other samples were mixed with 6x loading dye prior to be loaded on the gel. The voltage of 120-140 was applied for separation of the DNA fragments. To estimate the molecular weight of these fragments, the standard DNA marker (GeneRuler DNA Ladder Mix or GeneRuler 1 kb DNA Ladder) was used. For visualizing the DNA fragments, gels were stained in ethidium bromide bath for approximately 2 min. Background staining was removed by rinsing the gel in water bath before imagining. DNA bands were detected when the gel was exposed to ultraviolet light at a wavelength of 336 nm due to the large increase in fluorescence of the ethidium bromide upon binding to the DNA. Gels were documented by the Intas Gel iX Imager.

3.12.10 Extraction of DNA from agarose gels

For cloning procedures, DNA fragments and PCR should purify from the agarose gel prior to ligation. Therefore a small part of the gel containing the desired DNA fragment cut out of the gel by a clean scalpel. Subsequently the DNA fragment extracted from the gel using Qiagen gel extraction kit following manufacturers' instructions. The purified DNA was dissolved in 15-20 µl nuclease free.

3.12.11 λ-Red recombination

Phage lambda-derived Red recombination system is a powerful technique for genetic modification like generating insertion, deletion, and point mutations on chromosomal, plasmid, or BAC targets (Swaminathan et al., 2001; Thomason et al., 2007). Bacteria are not readily transformable with linear DNA due to their defense mechanism. The intracellular *recBCD* exonuclease degrades linear DNA. This problem can be solved using three phage-derived lambda Red proteins (Gam, Exo, and Beta) which are necessary for carrying out dsDNA recombination (Mosberg et al., 2010). Gam impedes the degradation of linear dsDNA by the *E. coli* RecBCD and SbcCD nucleases. Therefore the replacement of a DNA sequence with a construct of linear DNA can be facilitated (Sawitzke et al., 2007). These DNA fragments should contain a selectable marker as well as the homology arms at both ends. These homology extensions should be as short as 35–60 bp to targeted sequence. These homologous regions allow the integration of the linear DNA through the homologous recombination.

The selected host for performing λ-Red recombination was *E. coli* BW25113 harboring the helper plasmid pKD46. This plasmid carries the necessary λ-Red genes behind the *araBAD* promoter.

A single colony of *E. coli* BW25113 containing the siphonazole gene cluster, was inoculated in 3 ml LB medium supplied with Ampicillin (For pKD46) together with Apramycin (for pCC1FOS+siphonazole) and incubated overnight at 37°C. This culture was served as the pre-culture and was used for inoculation of 100 ml main culture. To induce the expression of recombination genes, *L*-arabinose was added to the main culture at the final concentration of 1 mM. The cells were incubated at 37°C until an OD_{600} of 0.5 was reached, and afterward were prepared for electroporation (3.11.2). These cells were transformed by the linear pESAC13 vector according to protocol 3.11.5.

The second cloning strategy using λ-Red recombination was done as following:

First, a single colony of *E. coli* BW25113 harboring pKD46 vector was inoculated in 3 ml LB medium containing Ampicillin and incubated overnight at 37°C. 1 ml of this culture was used for inoculation of 100 ml LB medium which then was incubated at 37°C until the OD_{600} ~0.5. The cells were prepared for electroporation and transferred via 5-7 µl of pESAC13 vector.

The obtained *E.coli* cell containing the pESAC13 vector plus pKD46 was inoculated in 3 ml LB medium containing 100 µg/ml of Ampicillin and Kanamycin (For pESAC13) and incubated overnight at 37°C. 1 ml of this culture was used for inoculation of 100 ml of LB medium. This culture was induced by adding 1 ml of a 1 M *L*-arabinose solution which then incubated at 37°C at constant shaking until OD_{600} was reached to 0.5. The grown cells were prepared to be serve as electro-competent cells and transformed by DNA fragments.

3.12.12 Heterologous expression of the putative DAHP synthase (Amb7)

The isolated genomic DNA of the *F.ambigua* was utilized as the template for amplification of the coding sequence by primer pair DAHP Fa Fwd and DAHP Fa Rev containing the sequences of the restriction sites of *Eco*RI and *Hin*dIII at 5' end of each primer. The resulting PCR product of 1146 bp was isolated from the agarose gel and ligated to the pGEM-T vector. The positive clones were picked, and the isolated plasmids were subjected to be sequenced. A accurate clone, according to restriction digestion pattern and sequencing, was selected and inoculated in LB liquid medium containing ampicilin for selection. The Isolated construct containing the *amb7* coding region was isolated, restricted at the introduced restriction sites with *Eco*RI and *Hin*dIII and ligated into the similarly restricted expression vector pET28. The constructed vector (pETDAHP Fa), as was verified by restriction digestion and amplification, was used to transform competent *E. coli* BL21 (DE3) cells. The latter strain served subsequently as expression host.

3.12.13 Heterologous expression of the didomain SphH Hyd-ACP

To amplify the coding sequence of the SphH Hyd-ACP, the primer pair SphH end fw and SphH end dn was designed. Forward primer contains the sequence of CACC, at its 5´ end which is identical to the GTGG sequence of pET TOPO® vector. pCC1FOS containing the complete siphonazole gene cluster was used as template for amplification of the target gene. Using the *pfu* DNA polymerase in the PCR reaction resulted in a blunt-ended fragment. The amplified fragment was extracted from the agarose gel when the correct size (1378 bp) of the fragment was affirmed by electrophoresis. The yielded DNA was cloned into the pET-TOPO® vector by topoisomerase cloning. The ready construct of pET151 harboring SphH Hyd-ACP was transferred to Top10 chemically competent *E. coli* cells. The grown colonies on LB medium supplied with ampicillin were picked and inoculated in 3 ml liquid medium. Subsequently, plasmids were isolated from these cultures (3.12.2) and subjected to sequencing. The primer pair T7 and T7 term was used for sequencing the expression construct. The accurate construct (pET151-SphH), according to sequence analysis, was transferred to the expression host *E. coli* BAP1.

3.12.14 Heterologous expression of the putative thioesterase, SphJ

For determination of the exact role of SphJ in siphonazole biosynthesis, this protein was heterologously expressed in *E. coli.* The primer pair, *SphJ*-topo-Up and *SphJ*-topo-Down was used to amplify the SphJ- coding sequence.

The amplified fragment (736 bp) was subsequently cloned in pET151and transformed to *E. coli* XL1-Blue. The positive colony containing the accurate construct with the exact nucleotide sequence of *sphJ*, as deduced form the restriction pattern and the sequencing result, was selected for further experiments. The resulting construct pET151_SphJ was isolated from the corresponding colony and used to transform the expression host *E. coli* BL21 star.

3.12.15 Heterologous expression of the loading module part of CorI

Genomic DNA of *C. coralloides* B035 was isolated with a Wizard Genomic DNA purification kit (Promega) according to the manufacturer's instructions. This DNA served as the PCR template for the amplification of the 5´ part of the *corI* gene to obtain an ACP1-containing fragment. Using a proofreading polymerase with the primer pair dn_ACP1_TOPO and up_ACP1_TOPO, a PCR product of 1,432 bp was amplified and subsequently purified by agarose gel chromatography. The DNA band was cut out and extracted using the Wizard SV gel and PCR clean-up system (Promega). This was introduced in the vector pET151 TOPO (Invitrogen) by topoisomerase cloning. Transformation of competent *E. coli* cells with this mixture followed, and ampicillin resistant clones were selected. The plasmid pET151_ACP1 was isolated and subsequently transferred into the expression host *E. coli* BAP1.

3.12.16 Heterologous expression of the SAM-dependent *O*-methyltransferase CorH

For the construction of the expression plasmid for *corH*, the CorH coding sequence was submitted to the online tool GeneOptimizer, which is embedded in the pipeline for gene synthesis by GeneArt (Invitrogen). The codon usage of the gene was optimized for *E. coli* as an expression host, and the recognition sites for the restriction enzymes *Eco*RI and *Hin*dIII were introduced upstream and downstream of the gene, respectively. Using these restriction sites, the gene was excised from the original plasmid, gel purified, and subsequently ligated into the likewise-restricted expression vector pET28, yielding plasmid pET28CorH*. The identity of the introduced CorH-coding sequence was verified by sequencing. We used *E. coli* BL21 as the expression host. The DNA sequence of the corallopyronin A biosynthetic gene cluster of *C. coralloides* B035 is available under GenBank accession number HM071004.

3.13 Molecular biological methods concerning proteins

3.13.1 Heterologous expression of proteins

For the *in vitro* investigation of proteins, an adequate amount of the targeted protein was necessary. Therefore, the DNA sequence of the desired protein was cloned into the expression vectors (pET28 or pET151). These expression vectors contain an isopropyl-β-D-thiogalactopyranosid (IPTG) inducible promoter as well as an affinity tag, 6-his tag, to facilitate the purification of the protein via affinity chromatography (3.13.3).

In the first step, the coding sequence of the targeted protein was amplified by PCR using *pfu*-polymerase and cloned to the expression vectors. The generated plasmid was first transferred to the *E.coli* XL1-blue where the colonies were screened to contain the correct construct. The proved construct was isolated and transferred to the *E.coli* BL21 or *E.coli* BAP1 cells, as the heterologous hosts.

A single colony of the corresponding *E.coli* cells containing the desired construct was inoculated in 10 ml LB medium supplied with the appropriate antibiotic. This culture was used for inoculation of 1 liter LB medium and inoculated mostly at 37°C by constant shaking. Protein expression was induced by adding IPTG to a final concentration of 0.4 – 0.5 mM and incubation was extended at 16°C over night.

3.13.2 Cell lysis by sonication

Cells from protein expression were harvested by centrifugation at 4°C and 4000 rpm for 20 min. To open the cells, pellet was resuspended in 20 ml of lysis buffer and the bacterial suspension was subjected to sonication using Branson Sonifier 250, set to output level 4.50 % duty cycle. The samples were sonified six times each for 10 pulses while they were placed in ice.

Cell debris and insoluble parts were pelleted by centrifugation at 4°C and at 8,500 rpm for 20 min. The supernatant containing the protein was collected and stored in ice. The

cell debris was re-suspended again in 10 ml lysis buffer and subjected to the next round of sonication and centrifugation for detection of insoluble fraction.

3.13.3 Purification of recombinant proteins by Ni-NTA- affinity chromatography

Ni-NTA Agarose is an affinity chromatography matrix for purifying recombinant proteins carrying a His tag. Histidine residues of the His tag bind to the vacant positions in the coordination sphere of the immobilized nickel ions with high specificity and affinity, while the other protein elute. Cleared cell lysate is loaded onto the matrix. During the washing and elution steps by increasing concentrations of imidazole unspecific linked proteins can be eluted due to the competition of imidazole with histidine for binding to Ni^{2+}.

His-tagged proteins are bound, and other proteins pass through the matrix. After washing, His-tagged proteins are eluted in buffer under native or denaturing conditions.

The twice positive charged Ni^{2+} ions interact with the histidine residues of the 6-his tag and bind the protein to the matrix while other proteins elute.

The Ni-NTA gravity flow column containing 1 ml Ni-NTA-agarose was prepared and equilibrated in lysis buffer. Afterward, the cleared lysate was loaded on the column and allowed to pass through the matrix. To increase the binding efficiency, the resulted flow-through was re-loaded on the column at least twice more. The column was washed twice with 2.5 ml washing buffers with 30 and 50 mM imidazole. Subsequent, the elution of the protein was performed in five elution steps using 500 µl elution buffer with 100, 150, 200, 300 and 300 mM imidazole, if not stated otherwise. All the collected fractions were kept in ice until being analyzed by SDS-Polyacrylamind gel electrophoresis.

3.13.4 Denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

This method is commonly used in the laboratory for the separation of proteins based on their molecular weight, under the influence of an applied electrical field. In the denaturizing approach, the fractions were boiled and treated with mercapto ethanol to

reduce all di-sulfide bonds and to gain unfolded proteins. During electrophoresis in a SDS milieu, the proteins are charged completely negative, thus allowing a strict separation by molecular weight.

The separating gel is formed by radical polymerisation of bis-acrylamide to polyacrylamide, which forms a molecular sieve. Depending on the protein size, different concentration of polyacrylamide can be chosen. For a better focusing of protein bands, discontinuous gels were used. The discontinuous buffer systems employ different buffer ions and pH in the gel and in the electrode reservoir. The reaction mixture is given below.

Once ammoniumperoxosulfate (APS) and N,N,N',N'-tetraethylendiamin (TEMED) is added, the gel will begin to polymerize. Thus the gel can be pipetted between the plates with a spacer distance of 1.0 or 1.5 mm and a layer of isopropanol covered the top of the gel solution. After polymerization, isopropanol was poured off using a filter paper and the stacking gel was added.

Glycin SDS-stacking gel (4%)

Tris/HCL pH 6.8 (1M)	375 µl
SDS (10%)	30 µl
Bis-acrylamide (30%)	510 µl
Water	2040 µl
APS (10%)	30 µl
TEMED	3 µl

Glycin SDS-separating gel (15%)

Tris/HCL pH 6.8 (1M)	2500 µl
SDS (10%)	100 µl

Bis-acrylamide (30%)	4000 µl
Water	3300 µl
APS (10%)	100 µl
TEMED	4 µl

For each run, the reservoirs of the electrophoresis assembly were filled with fresh SDS electrophoresis buffer. Protein samples were mixed in the ratio of 4:1 with the denaturing loading buffer and boiled for 10 min at 99°C. Samples as well as a molecular size marker were pipetted into the gel, adjusting the volume according to the amount of protein in the sample. Electrophoresis was performed in a XCell SureLock® Mini-Cell. The voltage was set at 90 V until the samples reached the separating gel, afterward it was increased to 130 V. After electrophoresis, the gels were analyzed by coomassie-staining

3.13.5 Coomassie-staining

Coomassie blue dyes are a family of dyes commonly used for visualization of proteins separated by SDS-PAGE. After electrophoresis, the gel was soaked in the staining solution and heated shortly in the microwave. The staining was continued for several minutes while it was shaking horizontally. The excess stain was eluted with the distaining solution within shaking for several hours or overnight. The gels were documented with the INTAS illuminator.

3.13.6 Concentration of purified proteins and buffer exchange

The samples which proved to contain the desired protein according to the SDS-PAGE result were concentrate using spin filter column (Millipore, 10 kDa exclusion size). This column was also used for exchanging the imidazole from the elution buffer with 50 mM

Tris-HCl (pH 7) buffer. All the fractions were combined and loaded to the column and centrifuged at 6000 rpm and 4°C until only one-tenth of the starting volume remained.

Afterward the column was loaded by Tris-HCl (pH 7) buffer with the same amount of the starting volume of the samples. Centrifugation was continued at the same condition until the final volume of 50-100 µl was achieved. The yielded protein was stored in the ice until used in the assay.

3.13.7 Determination of protein concentrations after Lambert-Beer

Since proteins absorb light at a specific wavelength, a spectrophotometer can be used for measuring the concentration of a protein in solution using direct UV measurement at 280 nm. The calculation is based on the Lambert-Beer equation (Formula 2-1). It requires the knowledge of the molar extinction coefficient ε, with the program ProtPram provided by the Swiss Institute for Bioinformatics. The UV absorbance of the protein samples in this study were measured using Thermo Scientific nanodrop.

Formula 2-1:

$$concentration: \left[\frac{mol}{L}\right] = \frac{OD \times d}{\varepsilon}$$

Division by molecular weight gives concentration in g / l.

OD_{280} = optical density at λ = 280 nm

d = dilution factor

ε = molar extinction coefficient [M^{-1} cm^{-1}]

3.14 Chromatography

3.14.1 Affinity chromatography

See Molecular biological methods concerning proteins (3.13.3)

3.14.2 High performance liquid chromatography (HPLC)

HPLC was used for detection of the attached substrate to the PPant arm of the ACP in the study of analysis of the unusual decarboxylation and dehydration in biosynthesis of siphonazole as well as ACP1 loading assay in biosznthesis of corallopyronin A.

HPLC was performed on a Merck-Hitachi system equipped with an L-6200A pump, an L-4500A photodiode array detector, a D-6000A interface with D-7000 HSM software and a Rheodyne 7725i injection system. Column was a waters symmetry 300 C4 (5 µm, 4.6 x 250 mm). The complete reaction mix was stopped by adding the equal amount of methanol and injected without further preparation and separated using a gradient of 0.1% TFA and acetonitrile as liquid phase. The complete setup is given below.

Solvent A: H_20 with 0.1% TFA

Solvent B: Acetonitrile

Gradient 0 min: 65% A, 35% B

45 min: 20% A, 80% B

50 min: 0% A, 100% B

Flow: 1 ml / min

3.14.3 Mass spectrometry

The LC-ESI MS was performed by E. Eguereva (Institute for Pharmaceutical Biology of the University of Bonn) employing an Agilent 1100 Series HPLC including DAD, with a RP 18 column (Macherey-Nagel Nucleodur 100, 125 mmx2 mm, 5 µm) coupled with an API 2000, Triple Quadrupole, LC/MS/MS, applied Biosystems/MDS Sciex and ESI source. The applied gradient elution of (From 90% H_2O to 100% MeOH in 10 min, then 100% MeOH to 20 min, with added NH4Ac, 2 nM, DAD 220.0–400.0 nm) was used for characterization and purity determination.

Analysis of the protein mass spectrometry was done by M. Sylvester (Institute for Biochemistry and Microbiology of the University of Bonn) using a Thermo LTQ Orbitrap Velo coupled with an Advion TriVersa NanoMate enabling a continues electron spray. A spray chip with 5 µm nozzle diameter was used at a spray voltage of 1.6 kV and 0.3 psi pressure setting.

3.14.4 Gas chromatography–mass spectrometry (GC-MS)

GC-MS is a technique for separation, identification and quantification of complex mixtures of chemicals by coupling gas chromatograph (GC) and mass spectrometer (MS).

This method was used for detection of the final off-loaded trans-1,3-pentadiene in analysis thioesterase domain of siphnazole biosynthesis gene cluster. The reaction mixtures were incubated for adequate period. Afterward, these samples were extracted with ethyl acetate; the upper phase was separated and analyzed for the presence of trans-1,3-pentadiene. 2-5 µl of the prepared samples was directly injected to the GC-MS and analysis was carried out on a The gas chromatography–mass spectrometry Perkin Elmer GCMS / TCD.

3.15 *In vitro* 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) assay

The activity of the DAHP synthase was measured using a continuous spectrophotometric method. The reaction mixture containing the freshly purified enzyme together with PEP (80 µM), E4P (350 µM), and $MnSO_4$ (100 µM) in 50 mM BTP buffer pH 7 was incubated in 1cm path length quartz cuvette and the dropping absorption was observed at 232 nm using LAMBDA 40 UV/Vis Spectrophotometer, PerkinElmer. One unit of enzyme activity is defined as consumption of 1 µmol of PEP per minute (Schofield et al., 2004). The production of the DAHP was proven by LC-ESI MS.

3.15.1 Metal requirement of DAHP synthase

To investigate the metal dependence of the DAHP synthase, the enzyme and the assay mixture were separately incubated with EDTA for 10 minutes prior the reaction. The reaction was started by adding EDTA treated Amb7 to the reaction mixture (80 µM PEP, 350 µM E4P, EDTA 100 µM, 50 mM BTP buffer; pH 7). Enzyme activity was monitored at room temperature. Furthermore, the effect of different metal ions on the catalytic activity of DAHP synthase was determined by substituting different divalent cations, i.e. Mg^{2+}, Zn^{2+}, Cu^{2+}, and Cd^{2+}, for Mn^{2+} (100 µM) in the standard reaction mixture.

3.15.2 Effect of temperature on DAHP activity

To determine the effect of temperature on the enzymatic activity of the DAHP-synthase, the reaction mixture containing 80 µM PEP, 350 µM E4P, 100 µM $MnSO_4$, and BTP buffer (50 mM, pH 7) was incubated for 10 minutes at the respective temperature between 10°C to 70°C. The reactions were started by the addition of enzyme.

3.15.3 pH dependence of DAHP synthase

To investigate the pH dependence of the DAHP-synthase, the pH of the BTP buffer used in the assay was varied (pH range: 5-9). Consumption of PEP was monitored at room temperature.

3.15.4 Determination of Kinetic properties

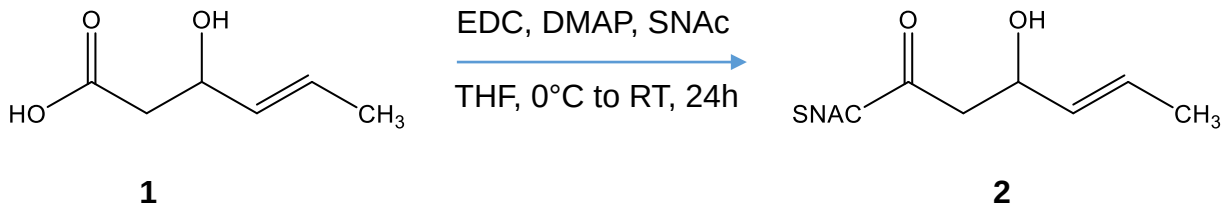
The beforehand determined optimum conditions were used for the kinetic studies of DAHP synthase. Thus, kinetic contents of Amb7 were determined according to double-reciprocal curves. The concentration of E4P was varied between 0.025 and 1 mM, whereas the concentration of PEP was kept constantly (80 µM). The reaction mixture contained 100 µM $MnCl_2$ were initiated by adding appropriate amount of the freshly

purified enzyme. The reaction was monitored at 30°C for several minutes. K_m, V_{max} and K_{cat} were determined based on the double- reciprocal plots.

3.16 *In vitro* assays regarding dehydration and decarboxylation

3.16.1 Chemical coupling of the model substrate

To analyze the unusual decarboxylation/release mechanism, a substrate mimic has to be synthesized. As starting point served the commercially available trans-3-hydroxyhex-4-enoic acid, which has to be activated by SNAC-coupling. The coupling reaction was done by Sarah Frank (Figure 3-1) (Institute for Organic Chemistry of the University of Bonn). Substrate **1** was purchased from Epsilun Chemie.



1 **2**

Figure 3-1. Coupling of substrate **1** with N-acetylcysteamine (SNAC) via 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), 4-dimethylaminopyridine (DMAP) in standard procedure gives substrate **2**.

3.16.2 phosphopantetheine (PPant) ejection assay for determination the fanction of putative Hydrolase part of SphH in biosynthesis of siphonazole

Activity of putative hydrolase (part of SphH) was determined in a MS based *in vitro* assay. This method uses the multi-stage tandem MS on a common ion trap instrument to obtain high-resolution measurements of the masses of substrates and intermediates bound to phosphopantetheinylated (*holo*) carrier proteins, which was successfully applied before (Meluzzi et al., 2008).

Thermal activation by infrared multiphoton dissociation (IRMPD) induces the PPant arm ejection of the *holo* form of acyl and peptidyl carrier proteins. This technique allows the mass of substrates loaded onto carrier proteins to be readily deduced from the mass of the corresponding PPant fragments (Hansen et al., 2007).

In the first step the SNAC bounded substrate should be loaded on the *holo* ACP. Afterward, analyzing the bounded molecule to the PPant arm can reveal the biochemical activity of the putative hydrolase.

Therefore, substrate **2** was loaded to the heterologously expressed didomain SphH Hyd-ACP by co incubation at room temperature as well as 30°C whitin different incubation periods. The reaction volume of 50 µl consisted of: 10 µL protein solution, 4 µL substrate, 36 µl buffer (Tris 50 mM, pH 7.0). The reaction mixture was incubated at room temperature or 30 °C for 1h, 2h or 16 h. To stop the reaction the equal amount of methanol was added to the reaction mixture and directly injected into the HPLC column (Waters Symmetry 300 C4; solvent: linear gradient from 65% water to 100% acetonitrile over 50 min, flow: 1.0 ml/min). The collected fractions were dried using Savant SpeedVac concentrator (AES2010 Centrifugal Evaporator). These samples were dissolved with the adequate volume of electro spray solution (49.5% H_2O, 49.5% methanol, 1% formic acid). 10 µL samples were loaded onto the 96 well plate of the NanoMate spray robot coupled to the LTQ Orbitrap Velos. A spray chip with 5 µm nozzle diameter was used at a spray voltage of 1.6 kV and 0.3 psi pressure setting. An environmental polysiloxane ion with *m/z* 445.12003 was used as lock mass for internal

calibration. Typical mass deviation was <2 ppm. Isolation and fragmentation were performed in the linear ion trap, detection of the final product spectrum was done with the Orbitrap analyzer.

3.16.3 Analyses of the ejected molecule by mass spectrometry for determination the function of putative TE domain in biosynthesis of siphonazole

To Analyze of the off-loaded reaction product, it has to be cleaved off the ACP. This can be done via two possibilities: First, chemically hydrolyzing the thio-ester while the substrate is attached to the PPant arm, by using KOH, second applying the catalytic activity of the thioesterase in the reaction mixture.

Therefore, to investigate the catalytic activity of the putative TE (SphJ), this domain was heterologously expressed and used in the assay.

In the first approach a reaction mixture with the total volume of 50 µl containing 10 µl protein solution (SphH Hyd-ACP), 4 µl substrate **2** and 36 µl buffer (Tris 50 mM, pH 7.0) was incubated for 1 h, 2 h or 16 h at room temperature or 30°C. Afterward the reaction mixture was treated by 1 mM KOH at 37°C for 10 min to release the free ACP-SH by hydrolysis. After incubation, the reaction was neutralized with 0.01 of 4 N HCL. For extraction of the free product, 50 µl of the Ethyl acetate was mix thoroughly with the mixture and centrifuged for 5 min at 13000 ×g. The organic phase containing the ejected product was either dried in Savant SpeedVac concentrator (AES2010 Centrifugal Evaporator) to be analyzed by mass spectrometry or directly injected to the Gas chromatography–mass spectrometry Perkin Elmer GCMS / TCD.

In the second approach, the reaction was done either in cooperation of two portions (SphH Hyd-ACP and SphJ) or just in presence of the SphJ. The reaction mixture with the total volume of 100 µl containing, 10 µl of each protein solution, 4 µl substrate **2** in buffer (Tris 50 mM, pH 7.0) was incubated for 1 h, 2 h or 16 h at room temperature or 30°C. The reaction was stopped by addition of 50 µl ethyl acetate and the mixture centrifuged for 5 min at 13.000 ×g. The organic phase of the mixture which supposed to

contain the off-loaded product was collected and analyzed by Gas chromatography–mass spectrometry Perkin Elmer GCMS / TCD as it described before (3.14.4).

As a reference trans-1,3-pentadiene was purchased from Sigma- Aldrich, this compound served as the standard in different reaction mixtures.

3.17 ACP1 loading assay

Purified ACP1 protein solution was supplemented with 100 µl of cell lysate (3.11.7). Cell lysate was directly used in the assay, and incubation of the cell lysate and the purified ACP1 was performed for 70 min at 30°C. To desalt and purify the protein sample, the assay mixture was directly subjected to semi preparative RPHPLC (a C4 reversed-phase column [Symmetry 300], 5 µM, 4.6 by 250 mm, Waters; eluent, water, 0.5% trifluoroacetic acid-acetonitrile; gradient, 70/30 to 35/65 over 30 min; flow rate, 1.0 ml/min). The protein fraction was manually collected and dried in a SpeedVac concentrator. The dried samples were dissolved in an adequate volume of electrospray solution (49.5% H_2O, 49.5% methanol, 1% formic acid). Next, 20 µl samples were loaded onto the 96-well plate of the NanoMate spray robot coupled to the LTQ Orbitrap Velos. A spray chip with a 5 µm nozzle diameter was used at a spray voltage of 1.6 kV and a 0.3-lb/in^2 pressure setting. An environmental polysiloxane ion with *m/z* 445.12003 was used as a lock mass for internal calibration. Isolation and fragmentation were performed in the linear ion trap; detection of the final product spectrum was done with the Orbitrap analyzer. The ejection ions were obtained by in source fragmentation, applying 50 to 65 V of fragmentation energy. The resulting ejection ions were isolated and subjected to further rounds of fragmentation to ensure that these ions were indeed phosphopantetheine (PPant) ejection ions tethered with substrate. A next round of fragmentation (MS^3) yielded the expected PPant arm (calculated, *m/z* 216.1267; measured, *m/z* 216.1277), which could be further fragmented (MS^4) and revealed the expected PPant signature ions (Quadri et al., 1998).

3.18 SAM-dependent *O*-methyltansferase activity assay

A typical assay consisted of 10 to 50 µl of enzyme solution, 10 µl of $MgCl_2$ (25 mM), 10 µl of substrate (100 mM), and Tris buffer (50 mM, pH 8.0) and was started by adding 1 µl of 50 mM SAM in a final volume of 100 µl. An assay without a substrate served as a negative control to determine the blank value. The assay was stopped by adding 100 µl of methanol and could be stored at 4°C. The sample was pelleted in a table-top centrifuge (15 min, 17,000 x *g*), and the supernatant was transferred to a new vial and dried in a SpeedVac concentrator. The analysis of the *O*-methyltransferase activity was performed by a coupled *S*-adenosine-homocysteine hydrolase assay as previously described (Schäberleet al.,2013) and was read out by a microplate reader (Sunrise-Basic Tecan; extinction measured at 412 nm). The values represent endpoint measurements after 26 h of incubation at 30°C.

4. Results & Discussion

4.1 In detail analysis of precursor formation in the biosynthesis of the ambigols

4.1.1 The putative ambigol biosynthetic gene cluster

The ambigols A-C (Figure 1-9) isolated from a filamentous cyanobacterial strain of *F.ambigua*, are highly chlorinated aromatic compounds, representing an unprecedented structural class (1.7). In the PhD thesis of M. El Omari the genome sequence data of *F.ambigua* were analyzed along with screening of a fosmid library, and the putative ambigol biosynthetic gene cluster was deduced. Using bioinformatic tools, it was shown that the ambigol gene cluster (14.9 kb) comprises nine genes encoding biosynthetic enzymes (Figure 4-1).

Figure 4-1. Putative biosynthesis gene cluster of ambigols, i.e. *amb*1-*amb*9. The functionalities encoded by the genes are given in table 4-1.

Table 4-1. The BLASTp search results of the encoded proteins involved in ambigol biosynthesis.

Protein	Size[a] (aa)/(bp)	Highest Homolgy[b]	Identity[c] (aa)
Amb1	483/1452	CrpH [*Nostoc* sp. ATCC 53789] Non-heme halogenase	283/485 (58%)
Amb2	487/1464	Hypothetical protein [*Fischerella* sp. PCC 9339]	213/487 (44%)
Amb3	358/1076	Hypothetical protein [*Fischerella* sp. PCC 9339]	135/339 (40%)
Amb4	204/615	Hypothetical protein [*Calothrix* sp. PCC 7103]	106/163 (65%)
Amb5	207/624	4-hydroxybenzoate synthetase [*Coleofasciculus chthonoplastes*]	113/183 (62%)
Amb6	511/1536	AMP-forming acyl-CoA synthetase [*Moorea producens*]	332/514 (65%)
Amb7	376/1131	phospho-2-dehydro-3-deoxyheptonate aldolase [*Myxosarcina* sp. GI1]	268/372 (72%)
Amb8	876/2631	Non-ribosomal peptide synthetase module [*Moorea producens*]	519/870 (60%)
Amb9	574/1725	FAD-dependent oxidoreductase [*Moorea producens*]	366/560 (65%)

The size of the corresponding genes is given in base pairs (bp), and the size of the proteins is given in amino acids (aa). [b] BLASTp results for the translated amino acid sequences from the ambigol biosynthetic gene cluster. Length of the amino acid sequence is shown in column two. [c] The numbers of amino acids (aa) identical to the highest homologue are given.

Two putative halogenases, Amb1 and Amb9, are encoded at the N-terminal and C-terminal of the ambigol gene cluster. Amb1 shows the highest identity toward the $FADH_2$-dependent halogenase CrpH involved in the biosynthesis of cryptophycins (Magarvey et al., 2006). Amb9 shares only 22 % identity to Amb1, however, both show the conserved motifs for $FADH_2$-dependent halogenases. Amb2 and Amb3 represent

CYP450 enzymes as exemplified in many natural product gene clusters (Walsh & Fischbach, 2010). These domains are postulated to construct the trimeric ambigol backbone by phenolic oxidative coupling performed on the carrier-bound substrate. Amb 5-7 are hypothesized to be involved in precursor formation and activation. Amb8 is the only encoded NRPS-like module in ambigols biosynthesis gene cluster. The protein shows the highest identity (60%) to a NRPS module of *Moorea producens* (accession number: WP_008188544).

Amb8 consists of a C domain followed by a PP binding site and a TE domain, but is lacking an A domain for activation and loading of an amino acid. The C domain is in this case not involved in peptide formation; rather it may be responsible for the transfer of 4-hydroxybenzoic acid-AMP to the PCP domain. On the PCP-tethered substrate halogenation to a 3-chloro-4-hydroxy-intermediate may take place, before the molecule is released by the action of the TE domain.

The current study focused on catalytical function of the encoded Amb7 in ambigol biosynthesis.

4.1.2 DAHP Synthase Amb7 and its involvement in the formation of the starter unit in ambigols biosynthesis

A BLASTp search with the amino acid sequence of Amb7 (Figure 4-1) confirmed the assignment of the enzyme to the DAHP synthase I family (pfam00793). Members of this family catalyze the first step of the biosynthesis of the aromatic amino acids phenylalanine, tyrosine, and tryptophan as well as of aromatic cofactors such as folate and quinones, in microorganisms and plants (Herrmann 1995). Figure 4-2 shows the primary function of DAHP synthase which catalyzes the condensation of PEP and E4P to yield the final product DAHP. In addition, another biological function of the enzyme is to regulate the amount of carbon that enters the shikimate pathway (Herrmann 2001).

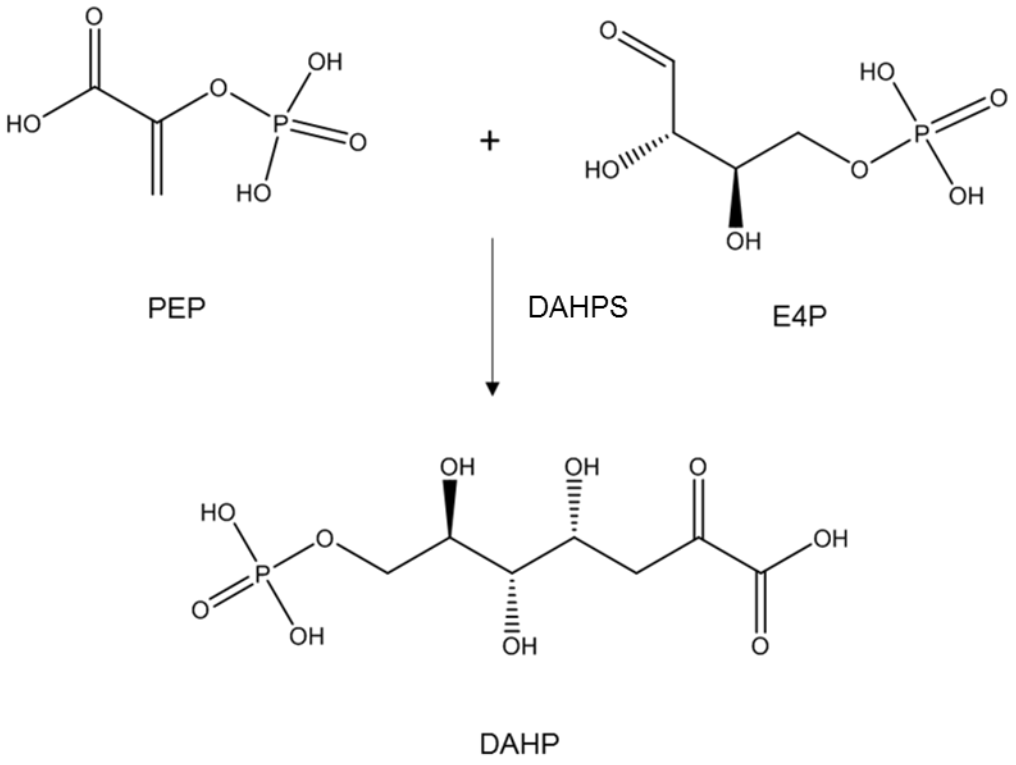


Figure 4-2. Biosynthesis of 3-deoxy-D-arabinoheptulosonate7-phosphate (DAHP) from phosphoenolpyruate (PEP) and erythrose-4-phosphate (E4P).

One striking feature of the DAHPSs is that despite their large sequence variability, all enzymes share a conserved tertiary core structure, a $(\beta/\alpha)_8$ TIM barrel fold. This core catalytic barrel is decorated with diverse small domains that are implicated in allosteric regulation of enzyme activity (Derrer et al., 2013; Wu & Woodard, 2006).

Multiple sequence alignment revealed the presence of the DAHPS signature motif D-x-x-H-x-N, as well as the highly conserved sequence motifs, K-P-R-T(S/T) and x-G-x-R in the Amb7 sequence (Figure 4-3).

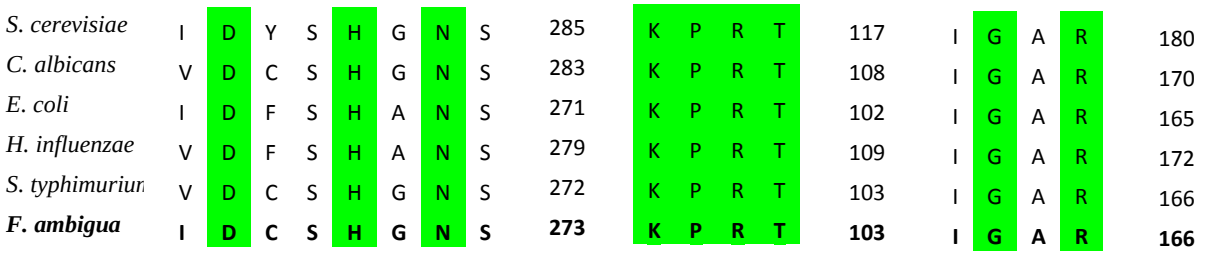


Figure 4-3. Multiple sequence alignment of the DAHP-synthase Amb7

The DAHPS signature motif D-X-X-H-X-N, K-P-R-T(S/T) and X-G-X-R are shaded in green. The reference sequences are taken from *Saccharomyces cerevisiae*, S288c852551 DAHPS (GI: 48425087), *Candida albicans* SC5314 (GI: 3647668), *Escherichia coli* K-12 (GI: 12932715), *Haemophilus influenzae* Rd KW20 (GI: 950411) and *Salmonella enterica* subsp. *enterica serovar Typhimurium* (GI: 1252280) (Gosset et al., 2001; Tapas et al., 2011).

Two unrelated families of DAHP synthase are classified as AroAI and AroAII. Smaller AroAI enzymes (30-40 kDa) are mainly of bacterial origin, while larger AroAII enzymes (>50 kDa) are more common in plants (Derrer et al., 2013; Subramaniam et al., 1998).

AroAI can be further divided in to two subfamilies: AroAIα and AroAIβ. According to former studies AroAIα and AroAIβ subfamilies differed by their metal requirements. Thus AroAIβ were proposed to be non-metallo enzymes, whereas the AroAIα enzymes are metallo enzymes. However, recent studies suggest that all DAHPSs are indeed metallo enzymes because of two reasons, i) some AroAIβ DAHPSs were clearly shown to require metals, and ii) the presence of an absolutely conserved metal binding sequence motif in all DAHPSs were identified (Wu & Woodard, 2003; Wu & Woodard, 2006; Birck & Woodard, 2001).

Crystal structures suggest that the active site metal binds to the four same residues in an octahedral geometry, i.e. in the *Aquifex aeolicus* KDO8P synthase (Cys^{11}, His^{185}, Glu^{222}, and Asp^{233} ligands to Cd^{2+}) as well as in the Phe-sensitive DAHP synthase of *E. coli* (Cys^{61}, His^{268}, Glu^{302}, and Asp^{326} ligands to Pb^{2+} or Mn^{2+}) (Wu et al., 2003; Igor A. Shumilin et al., 2002; Duewel et al., 2001). These described metal binding active site residues are absolutely conserved in Amb7 (Cys^{62}, His^{270}, Glu^{305}, and Asp^{330}) (Figure 4-4).

S. cerevisiae	G	P	C	S	I	78	S	H	G	N	284
C. albicans	G	P	C	S	I	72	S	H	G	N	282
E. coli	G	P	C	S	I	63	S	H	A	N	270
H. influenzae	G	P	C	S	I	70	S	H	A	N	278
S. typhimurium	G	P	C	S	I	64	S	H	G	N	271
F. ambigua	**G**	**P**	**C**	**S**	**I**	**64**	**S**	**H**	**G**	**N**	**272**

S. cerevisiae	I	E	S	N	318	T	D	A	C	344
C. albicans	I	E	S	N	316	T	D	A	C	345
E. coli	V	E	S	H	304	T	D	A	C	328
H. influenzae	V	E	S	H	312	T	D	A	C	336
S. typhimurium	I	E	S	N	305	T	D	A	C	330
F. ambigua	**L**	**E**	**S**	**N**	**306**	**T**	**D**	**K**	**C**	**331**

Figure 4-4. Multiple sequence alignment of the DAHP synthase Amb7 (metal binding residues).

Putative metal binding residues are shaded in yellow (Cys^{61}, His^{268}, Glu^{302}, and Asp^{326} based on amino acid sequence of the *E. coli* DAHP synthase. The reference sequences are taken from *Saccharomyces cerevisiae*, S288c852551 DAHPS (GI: 48425087), *Candida albicans* SC5314 (GI: 3647668), *Escherichia coli* K-12 (GI: 12932715), *Haemophilus influenzae* Rd KW20 (GI: 950411) and *Salmonella enterica* subsp. *enterica serovar Typhimurium* (GI:1252280) (I A Shumilin et al., 1999; Wagner et al., 2000).

4.1.3 Heterologous expression and purification of the putative DAHP synthase (Amb7) and enzyme assay

To investigate the role of Amb7 in ambigol biosynthesis, heterologous expression of this protein was projected. Amb7 was hypothesized to be involved in providing enough precursors for the shikimate pathway by catalyzing the first reaction in the mentioned pathway. The yielded product of the shikimate pathway can be further used by Amb5 to produce 4-hydroxy benzoic acid (4-HBA), the first precursor of ambigol biosynthesis. Therefore, the coding sequence of the Amb7 was cloned into the expression vector pET28 which resulted in the (pETDAHP Fa) construct (3.12.12). Afterwards, the constructed vector was transfer to *E. coli* BL21 (DE3) cells. The latter strain served subsequently as expression host.

Amb7 was overexpressed and purified based on its attached his-tag, via affinity chromatography on Ni-NTA column (Methods 3.13.1 & 3.13.3). Denaturing SDS-PAGE analysis (Figure 4-5) showed a prominent band in the range of 45 kDA. Elution fractions 1–5 were pooled, concentrated and re-buffered in Tris HCl buffer (pH 7). The protein concentration was determined to be around 1mg/ml, using the UV measurement at 280 nm with Thermo Scientific NanoDrop.

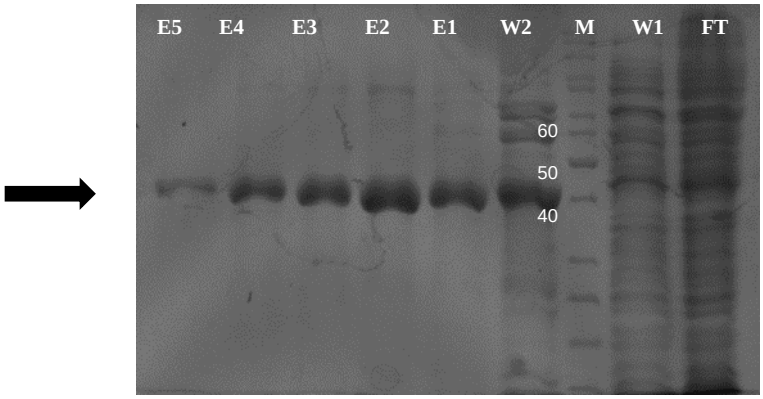


Figure 4-5. Purification of DAHP synthase Amb7 by affinity-chromatography.

SDS PAGE of a typical purification of Amb7. FT, flow through; W1 and W2, (washing steps with 30 and 50 mM imidazole, respectively); E1–E5, elution fractions (100-350 mM imidazole); M, marker (kDa). Black arrow indicates calculated mass of 45 kDA for the Amb7.

The purified protein was always utilized directly after expression and purification in the enzyme reaction mixture. A standard reaction mixture containing PEP, E4P, $MnSO_4$ and BTP buffer was incubated with the Amb7 to prove its activity. Consumption of PEP in the reaction mixture was determined by a continuous spectrophotometric measurement (Method 3.15). Dropping in absorption at 232 nm in the mixture containing Amb7, indicated enzymatic activity, while no difference in absorption was observed in the negative control, i.e. in the mixture without enzyme. Using this experimental design Amb7 activity, including the optimum temperature and pH were analyzed. In addition metal requirements, as well as kinetic constants of Amb7 were obtained (See 4.1.6 and 4.1.7).

4.1.4 Effect of temperature of the enzymatic activity of Amb7

Enzymatic activity of Amb7 was determined at different temperatures in the range of 10°C to 70°C. The assay mixture was pre-incubated for 5 minutes at the desired

temperature, and the reaction was started by adding the enzyme. Subsequently, the activity of the enzyme was measured by following UV absorption. The highest enzyme activity was shown to take place between 23°C and 40°C (23°C: 81.25±6.25%, 30°C: 100±0% and 40°C: 75±0%). Then the optimum activity was observed at 30°C. No activity was observed at 50-70°C, and just slight activity was presented at 10°C (12.5±0%) (Figure 4-6).

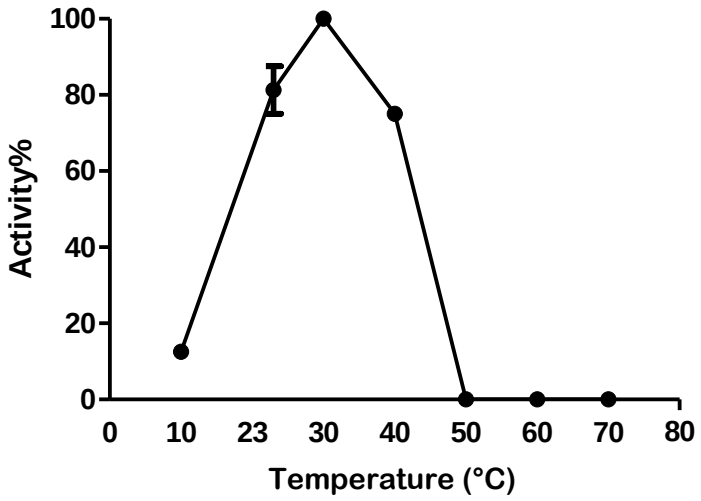


Figure 4-6. Effect of different incubation temperatures on the enzymatic activity of Amb7. Reaction mixtures containing 80 µM PEP, 350 µM E4P, 100 µM $MnCl_2$, as well as BTP buffer, were incubated at different temperatures for 5 minutes and the reaction was started by adding Amb7 (Method 3.15). UV absorptions were recorded 5 minutes after the onset of the reaction. Optimal activity was set to 100 %, whereby 100% activity are equivalent to a consumption of 80 µM PEP in 5 minutes. Measurements were performed in duplicate, the mean value of two independent protein purifications is given (10°C: 0±0%, 23°C: 81.25±6.25%, 30°C: 100±0%, 40°C: 75±0%, 50-70°C 0±0%).

4.1.5 pH dependence of the enzymatic activity of Amb7

To determine the effect of pH on Amb7 activity, the reaction was monitored at different pH values, i.e. between pH 5 and pH 9. Amb7 showed the best activity in the range from pH 6 to pH 8, whereby optimal activity, i.e. 100±0% was observed at pH 7. 100 % activity was calculated as the consumption of 80 µM PEP during 5 minutes.

At pH 5 Amb7 is still highly active, that is around 70% of the optimal activity (70.82±4.18%). At pH 9 the enzyme activity significantly dropped (35.35±2.15%) (Figure 4-7).

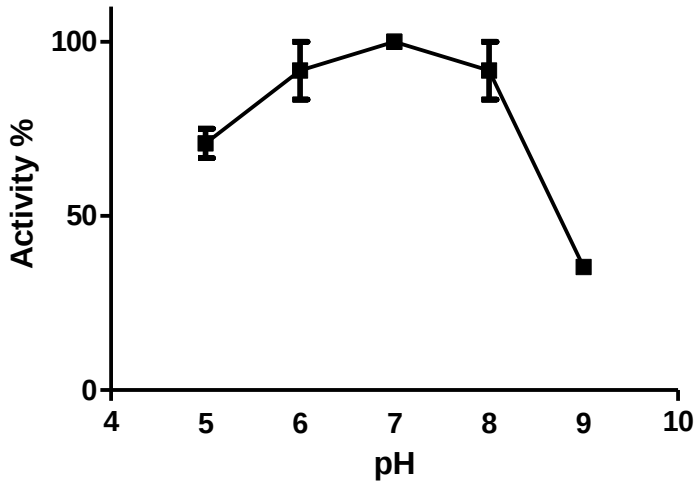


Figure 4-7. pH dependence of the enzymatic activity of Amb7.

Disappearance of PEP in the reaction mixture containing 80 µM PEP, 350 µM E4P, 100 µM $MnSO_4$ and BTP buffer (pH 5-9) was observed by adding Amb7, at room temperature. Values were recorded 5 minutes after the onset of the reaction (Method 3.15). 100% activity are equivalent to a consumption of 80 µM PEP in 5 minutes. Error bars correspond to the standard deviation of two determinations. The mean value resulted from two independent protein purifications is given, i.e. (pH 5: 70.82±4.18%, pH 6: 91.67±8.33%, pH 7: 100±0 %, pH 8: 91.67±8.33 and pH 9: 35.35±2.15%).

4.1.6 Metal requirement of DAHP synthase and effect of a metal chelating agent

The effect of divalent metal ions on Amb7 activity was analyzed. Therefore, different metal ions, i.e. Mn^{2+}, Mg^{2+}, Zn^{2+}, Cu^{2+} and Cd^{2+} were added to the standard reaction mixtures. The results indicated that utilizing Mn^{2+}, Zn^{2+} and Mg^{2+} increased enzyme activity, whereas incubation of the enzyme with, $Cd^{2+}$and Cu^{2+} resulted in lower activities (Table 4-2).

In a negative control experiment, the effect of the metal chelator EDTA was investigated in this study. As shown in table 4-2, DAHP synthase was inactive when EDTA was added to the reaction mixture.

Table 4-2. The effect of different metal salts and EDTA on Amb7 activity.

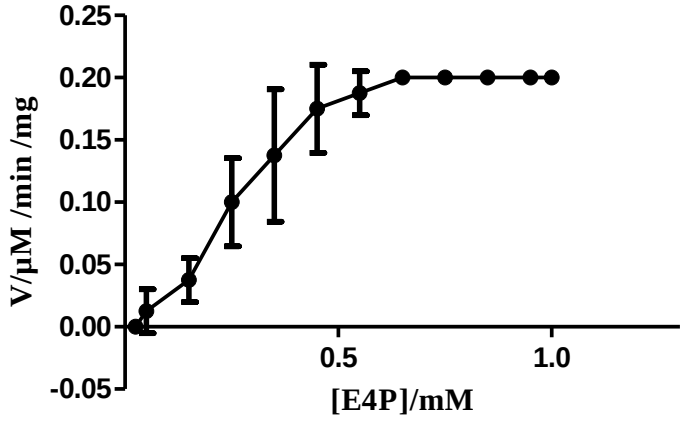
Metal ions	[a]Activity (unit)
$MnCl_2$	7.5±0.5
$MgCl_2$	7.0±1.0
$ZnCl_2$	7.5±0.5
$CdSO_4$	5.0±1
$CuSO_4$	2.0±1.0
EDTA	0±0

[a]One unit of enzyme activity is defined as the consumption of 1 µmol of PEP per minute. To investigate the metal dependence of the DAHP synthase, the enzyme and the assay mixture were separately incubated with EDTA for 10 minutes prior reaction. The reaction was started by adding EDTA treated Amb7 to the reaction mixture (80 µM PEP, 350 µM E4P, EDTA 100 µM, 50 mM BTP buffer; pH 7). Determination of metal ions on Amb7 activity was obtained in the reaction mixtures by substituting different divalent cations, i.e. Mg^{2+}, Zn^{2+}, Cu^{2+}, $Cd^{2+}$and Mn^{2+} (100 µM). Assay was performed as

described in method 3.15. The mean value of two independent protein purifications obtained after 10 minutes of incubation of the reaction mixtures with Amb7, are given in table 4-2.

4.1.7 Kinetic properties of Amb7

After the determination of the metal requirement, optimal pH and temperature, the kinetic properties of Amb7 were identified. Thus, using Graf Pad Prism 5, kinetic constants of Amb7 were determined according to double-reciprocal curves (Figure 4-8). Double-reciprocal plot of the Amb7 was calculated against the substrate concentration. The concentration of PEP was always preserved at 80 μM, while the concentration of E4P was varied between (0.025-1.000 mM). The maximum velocity achieved by the system, at saturating substrate concentrations (V_{max}) was determined to be 0.3725±0.083 (U/mg^{-1}) and the obtained K_m and K_{cat} values were shown to be 0.6993± 0.3746 mM and 0.0281±0.063 s^{-1}, respectively.



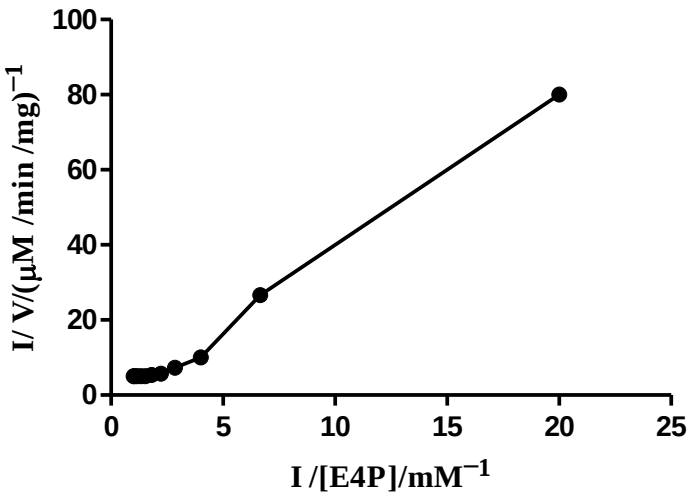


Figure 4-8. Lineweaver-Burk plot of DAHP synthase Amb7 reaction rate against substrate concentration. The erythrose-4-phosphate concentrations was varied (0.025-1.000 mM) while phosphoenolpyruvate concentration was maintained at 0.08 mM. The reaction, containing 100 µM $MnSO_4$, was initiated by the addition of Amb7. Reaction systems were incubated for 10 minutes at room temperature. The UV absorption at 232 nm was recorded after 10 minutes of the adding Amb7 for each reaction. The mean value of two independent protein purifications are given.

4.1.8 Discussion: DAHPS Amb7 and the origin of the starter unit in ambigols biosynthesis

The three following encoded proteins Amb5, Amb6 and Amb7 are hypothesized to take part in precursor formation for ambigols (El Omari. PhD thesis). Amb5 showed highest identity (62%) to the 4-HBA synthase from *Coleofasciculus chthonoplastes.* Amb6 had 65% identity to an AMP-forming acyl-CoA synthase from *Moorea producens,* and Amb7 is a DAHPS. It was assumed that the obtained chorismate from the shikimat pathway leads to 4-HBA by elimination of pyruvate. This can be catalyzed by putative 4-HBA

synthase (Amb5). Afterward, 4-HBA can subsequently be activated by Amb6 which catalyses the adenylation of 4-HBA to 4-HBA-AMP.

The conversion of chorismate to 4-HBA was determined in previous studies. For example the diffusible factor synthase XanB2 from *Xanthomonas campestris* pv catalyzes the hydrolysis of chorismate to 3-HBA and 4-HBA. 3-HBA and 4-HBA, are involved in the biosynthesis of the yellow pigment, xanthomonadin and antioxidant activity in *X. campestris* pv, respectively (Zhou et al., 2013).

Also in Gram-negative *E. coli* the conversion of chorimate to 4-HBA was well observed. In analogy to the proposed mechanism for Amb5, the chorismate lyase UbiC catalyzes the production of 4-HBA directly from chorismate by elimination of the enol-pyruvyl side chain. Heterologous expression of this protein in chloroplasts of plants which do not contain an orthologue of this protein resulted in production of large amount of 4-HBA (Alt et al., 2011).

4-HBA is assumed to be the basic monomer of the ambigols. Figure 4-9 shows the hypothetical scheme of precursor formation in ambigol biosynthesis. The DAHP synthase Amb7, encoded in the ambigol biosynthetic gene cluster, was here investigated in detail in *in vitro* assays. This enzyme is providing enough precursors for the shikimat pathway by the interconnection of PEP and E4P. Finally the synthesized 4-HBA subsequently can be activated by Amb6 which catalyzes the adenylation of 4-HBA to yield 4-HBA-AMP.

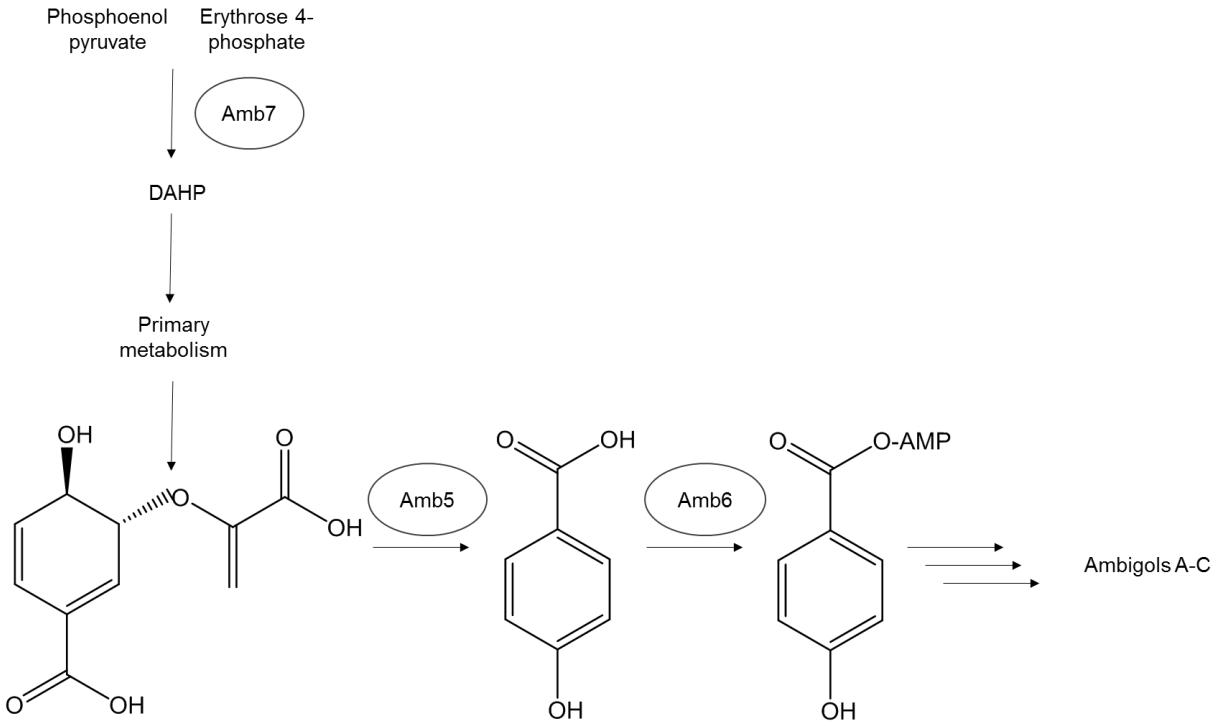


Figure 4-9. Hypothesis on starter unit formation in ambigol biosynthesis. The proposed role of DAHP synthase (Amb7), chorismate lyase (Amb5) and AMP synthase (Amb6) in 4-HBA-AMP formation.

It is proposed that Amb7 is a DAHPS isoenzyme, and involved in regulatory ambigol pathway. The presence of such isoenzymes like Amb7 in gene clusters corresponding to secondary metabolism is exemplified in many cases, e.g. in rifamycin and balhimycin biosynthetic pathways (Heinz G. Floss & Yu, 1999; Shawky et al., 2007). In a metabolic engineering approach for increasing the precursor supply for balhimycin production, the role of Dahp as limiting factor for high-yield glycopeptide production was shown. Whereby, expression of an additional copy of the *dahp* gene from the balhimycin biosynthetic gene cluster resulted in improved specific glycopeptide production, while the deletion of *dahp* presented significant reduction in balhimycin production (Thykaer et al., 2010). The regulation of these isoenzymes is usually different than the one from

primary metabolism, the latter providing essential amino acids. In that way the precursor supply for the secondary metabolites is ensured.

The catalytical function of the DAHP synthase Amb7 was investigated in this study. Thus, the optimal conditions for the activity of the enzyme, e.g. metal requirement, pH dependence, and optimal temperature were examined.

Several DAHP synthases of different organisms have been described to be EDTA sensitive and thus require bivalent metal ions (Schnappauf et al., 1998). To gain further insight into the role of the metal ion and the catalytic mechanism in general, the crystal structures of DAHP synthase (Aro4p) from *S.cerevisiae* and different metal ions and ligands have been determined by König and co-workers. The crystal structure of Aro4p provides evidence that the simultaneous presence of a metal ion and PEP results in an ordering of the protein into a conformation that is prepared for binding the second substrate E4P (König et al., 2004). We could show that the enzyme Amb7 requires metal ions for being active, since treatment of the enzyme with EDTA resulted in complete loss of activity. This was expected since sequence alignment showed all the four metal binding residues are absolutely conserved in Amb7. The highest Amb7 activity was observed in presence of Mn^{2+}, Mg^{2+} and Zn^{2+} (Table 4-2). Similarly, Wu and co-workers observed that *apo*-DAHP synthase from *Thermotoga maritima* gained its activity by a variety of divalent metals. Incubation of the *apo*-DAHP with Mn^{2+} and Zn^{2+} resulted in the highest activity of this enzyme, as compared to the other metal ions (Wu et al., 2003).

The optimal activity of Amb7 was observed at pH 7, while at pH 9 the enzyme activity significantly dropped. The highest enzyme activity was obtained at 30°C. Amb7 was active between 23°C and 40°C, but no activity was observed at higher temperatures.

The optimum, neutral pH and relatively low temperature of Amb7 reflect the growth conditions of the original host *F. ambigua* which was isolated from shallow water near Mellingen in Switzerland. This was analogous to the DAHPS (dHplP) from *S. cerevisiae* which shows a pH optimum for its activity around 6.8 as well as DAHPS from *T.*

maritima with the optimum of pH 6.3 (Paravicini et al., 1989; Wu et al., 2003).A key influencing factor on the catalytic activity of an enzyme is the amount of available substrates (Purich, 2010). Michaelis-Menten kinetics can be used for quantifying the affinity of an enzyme to its substrates. In this study the Michaelis-Menten constants (K_m), maximal initial velocity (V_{max}), and K_{cat} values of the Amb7 were determined for the substrate E4P (Figure 4-8). The maximal velocity, or V_{max}, reflects how fast the enzyme can catalyze the reaction. This kinetic parameters of Amb7 are in the same range as other DAHP synthases, e.g. maximum velocity of Amb7 (V_{max} 0.3725±0.083 U/mg^{-1}) is comparable to that of NCgl0950 DAHP synthase from *C. glutamicum* ATCC 13032 (V_{max} 0.46±0.02 U/mg^{-1}) (Ma et al., 2012). The K_m value of Amb7 (0.6993± 0.3746 mM) was much more higher than that of NCgl 0950 DAHP synthase of *C. glutamicum* (0.29±0.03), which indicates that NCgl 0950 has higher affinity to E4P. In contrast DAHP synthase from *S.cerevisiae* with the higher K_m value of 0.5000 mM, shows lower affinity to E4P compare to Amb7 (Alt et al., 2011).

4.2 Introduction to the putative siphonazole biosynthesis gene cluster

Siphonazole (Figure 1-12) is the first secondary metabolite from *Herpetosiphon* sp. 060 that has been characterized (Nett et al. 2006). Further, it shows promising antiplasmodial activity (Nett. PhD thesis). The origin of all building-blocks of the molecule was elucidated by feeding experiments with labeled precursors (Nett. PhD thesis). By analyzing the genome sequence data, along with the results of the feeding experiment the putative siphonazole biosynthetic gene cluster could be deduced (Figure 4-10) (Höver. PhD thesis).

The siphonazole gene cluster covers approximately 50 kb, comprises 10 open reading frames and encodes for a PKS/NRPS-hybrid multi enzyme complex. The genes were consequently denominated *sphA*–*sphJ*.

Table 4-3 shows the *in silico* predicted functions of the proteins encoded in the siphonazole gene locus.

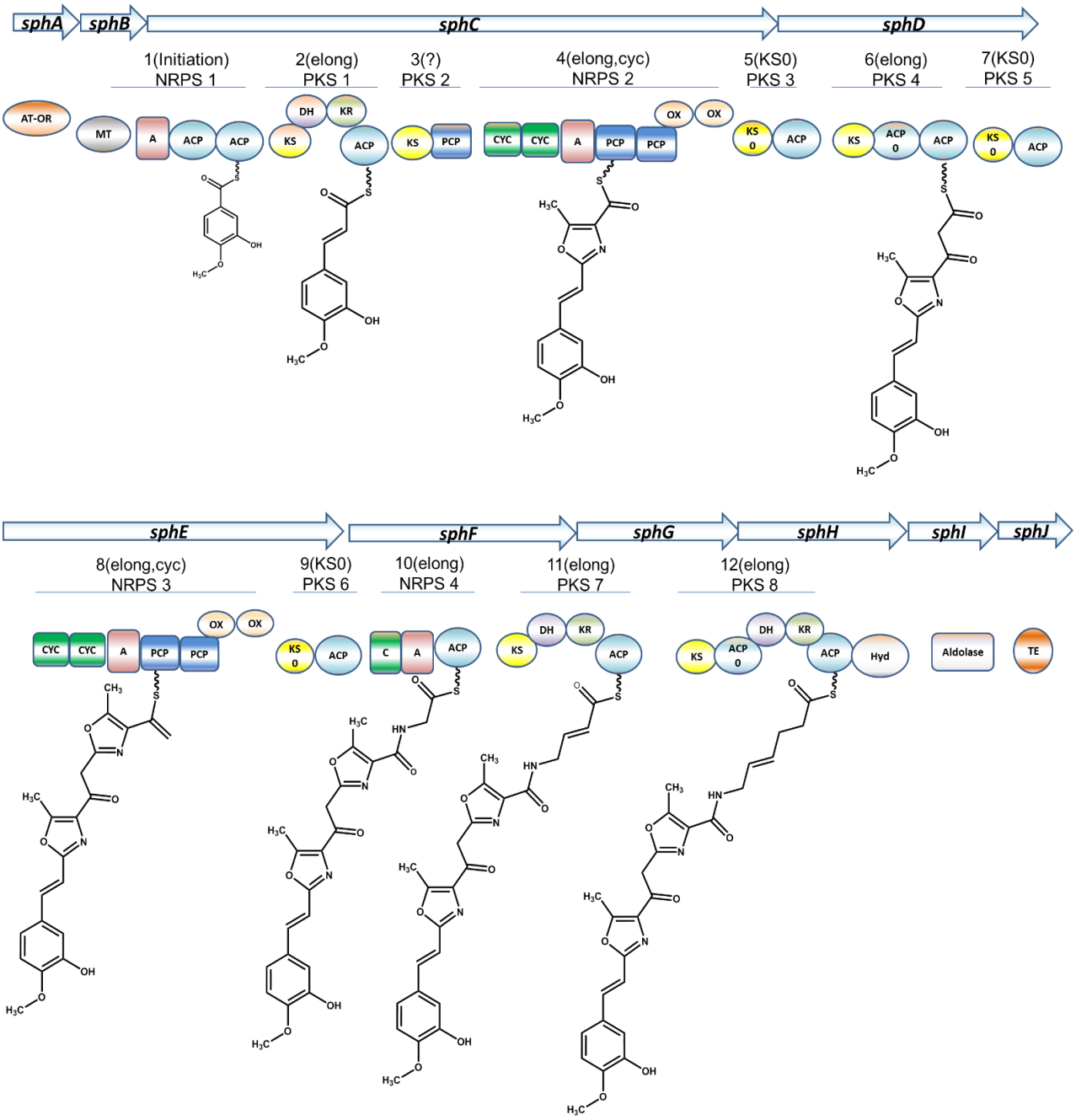


Figure 4-10. Putative pathway for the biosynthesis of siphonazole, A, adenylation domain; ACP, acyl carrier protein; Cyc, cyclisation domain; C, condensation domain; DH, dehydratase; Hyd, hydrolase; KR, ketoreductase; KS, ketosynthase; KS0, non-elongating KS; MT, methyltransferase; OR, oxidoreductase; Ox, oxidation domain; PCP, peptidyl carrier protein; TE, thioesterase.

Table 4-3. BLASTp search results for the translated amino acid sequences from *sph* genes.

Gene	Size[a] (aa)	Highest Homology[b]	Identity[c] (aa)	Predicted domains
sphA	762	malonyl CoA-ACP transacylase [*Pelosinus fermentans*]	389/765 (51%)	*trans*-AT-OR
sphB	221	SAM-dependent methyltransferase [*Paenibacillus sonchi*]	144/220 (65%)	*O*-methyltransferase
sphC	5451	Hypothetical protein [*Fischerella* sp. PCC 9339]	135/339 (40%)	A-ACP-ACP-KS-DH-KR-ACP-KS-PCP-Cyc-Cyc- A-PCP-PCP-Ox-Ox-KS
sphD	2089	beta-ketoacyl synthase [*Clostridium cellulolyticum*]	719/1951 (37%)	ACP-KS-ACP-ACP-KS-ACP
sphE	2693	hypothetical protein [*Paenibacillus polymyxa*]	662/2120 (31%)	Cyc-Cyc-A-PCP-PCP-Ox-Ox-KS-ACP
sphF	2059	mixed PKS/NRPS, partial [*Streptomyces avermitilis*]	862/2013 (43%)	C-A-PCP-KS-DH
sphG	1186	hypothetical protein, partial [*Clostridium cellulolyticum*]	479/1031 (46%)	KR-ACP-KS-ACP
sphH	1228	putative Carboxyl esterase [*Xenorhabdus bovienii* str. *oregonense*]	270/915 (30%)	DH-KR-ACP-Hydrolase
sphI	348	hypothetical protein [*Atribacteria bacterium* JGI 0000059-I14]	183/342 (54%)	Aldolase
sphJ	243	MULTISPECIES: thioesterase [*Bacillus*]	100/231 (43%)	TE

[a] The size of the proteins is given in amino acids (aa). [b] BLASTp results for the amino acid sequences from the siphonazole biosynthetic gene cluster. [c] The numbers of amino acids (aa) identical to the highest homologue are given; column five shows the predicted protein functions and NRPS and PKS domains according to the predictions derived from CLUSEAN and BLASTp search (Höver. PhD thesis).

4.2.1 Heterologous expression of the complete gene cluster

To prove the correspondence of the biosynthetic gene cluster to siphonazole production, first knock-out-experiments were envisaged. However, even by applying different transformation methods, e.g. heat shock, electroporation, sonoporation, conjugation and chemical transformation of protoplasts, no positive results were obtained. The strain *Herpetosiphon* sp. 060 was not accessible for genetic work (Höver. PhD thesis). Thus, heterologous expression of the complete siphonazole gene cluster was programed. Heterologous expression of the genes responsible for biosynthesis in a tractable host organism is one of the most useful strategies for advancing the study of natural products (Jones et al., 2013).

In a previous study by T. Höver, a fosmid (pCC1FOS-based) construct with a size of 71 kb, carrying the complete sequence information from three overlapping fosmids (EC9, EC10 and IC2), was constructed by applying the λ-Red recombination techniques. Since the heterologous expression of the gene cluster in *E.coli* resulted in no production of siphonazole, other hosts like Bacilli, as well as in Streptomycetes were candidates for further experiments. Bacilli were chosen according to the most similar codon usage patterns and Streptomycetes for their proven potential for producing a variety of secondary metabolites (Gomez-Escribano & Bibb, 2011).

4.2.1.1 Heterologous expression of the complete gene cluster in *Bacillus amyloliquefaciens* FZb42

The first attempt for heterologous expression of siphonazole was programed to be done in *Bacillus amyloliquefaciens* FZb42. To integrate the construct harboring the complete biosynthetic gene cluster into the *Bacillus* genome, the gene *baeJ* should serve as target sequence for homologous recombination. *baeJ* encodes the BaeJ protein which has the molecular functions of phosphopantetheine binding and transferase activity in *B. amyloliquefaciens*. As the proof of principle a smaller vector containing *baeJ* should be integrated into the *B. amyloliquefaciens* genome. Therefore, the gene was amplified and ligated into the vector pGF27 which contains the zeocin resistance gene (Zeo®). Transformation of *B. amyloliquefaciens* with the constructed vector (pGF27+*baeJ*) was performed by electroporation according to protocol 3.11.6.

The therefrom resulting colonies on LB medium supplied with 25 µg/ml Zeocin® were picked and the genomic DNA of these colonies was isolated according to protocol 3.12.1. A primer pair (Großer Pcil-R and Kleiner Asel-F) which amplifies the zeocin resistance gene was used to test for the presence of the vector-derived gene in the genome. The presence of this amplificate (~2500 bp) was proven in three independent colonies (Figure 4-11).

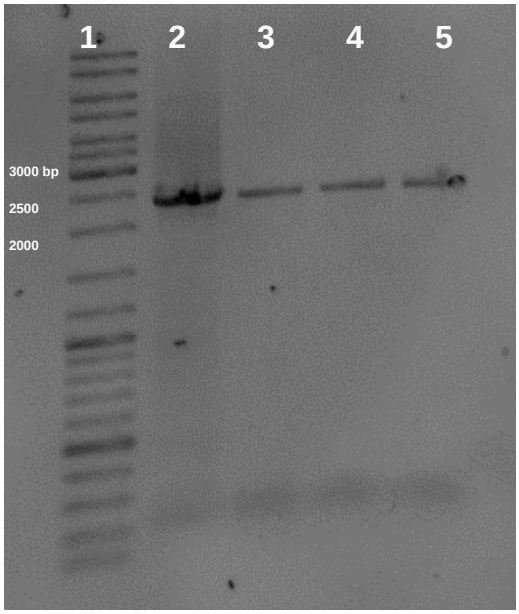


Figure 4-11. Analytical agarose gel of a zeocin resistance gene (~2500 bp).

1: DNA-ladder; 2: plasmid DNA (pGF27+*baeJ*) as positive control; 3-5: independent *E.coli* clones containing constructed vector pGF27+*baeJ*. The amplification of the fragment using genomic DNA as template proved the successful transfer of pGF27+*baeJ* into *B. amyloliquefaciens* FZb42.

By establishing the suitable process of transforming *B. amyloliquefaciens* FZb42 with the smaller vector the second step was to apply the same conditions to transfer the vector containing the complete siphonazole gene cluster as well as the *baeJ* gene.

Since in the last study (Höver. PhD thesis) the λ-Red recombination method was shown to be an appropriate method for rebuilding of the siphonazole gene cluster, it was planned to use this method for the construction of a vector containing the cluster and the *baeJ* gene for integration.

Most of the bacteria are not readily transformable with linear DNA due to the degradation of the transformed linear DNA as a part of bacterial defense mechanism by *recBCD* exonuclease (Exo V). To overcome the intracellular exonucleases, the targeted strain should harbor the Redα (*exo*), Redβ (*bet*), and Redγ (*gam*) proteins of the phage λ (Gust et al., 2003; Y. Zhang et al., 2000). In this study *E. coli* BW 25113 harboring the

helper plasmid pKD46, which carries the necessary λ-Red (*gam*, *bet*, *exo*) recombination functions, was used as the host for recombination.

The EC10+*bae*j vector containing the upstream part of the siphonazole gene cluster (starts ~5.5 kb upstream of *sphA* and ends within the sequences encoding the Cyc domains of *sphE*) as well as zeocin resistance gene, obtained from the previous study, was used as the template for the following PCR reaction. Primer pairs pGF27 +EC10 lamda-r and pGF27 +EC10 lamda-f were applied for amplification of the *baej*-Zeo fragment. Therefore, a 4 kb fragment containing the *baeJ* gene, Zeocin® resistance gene as well as homology arms at both ends to the nucleotide sequence of the EC10 vector was yielded.

The linear fragment obtained by PCR was purified and transferred to *E.coli* BW25113 containing the helper plasmid, pKD46 as well as the fosmid carrying the complete siphonazole gene cluster (pCC1FOS-siphonazole) by electroporation (3.11.5). The fosmids of the colonies obtained by this experiment were isolated for further investigation. To test if the fosmids carry the *baeJ* gene, necessary for recombination in *Bacillus*, a test-PCR with the primer pair *baej* fwd and *baej* rev was performed. The desired clones should result in the amplification of the *baej* fragment with expected size of 1243 bp (Figure 4-12).

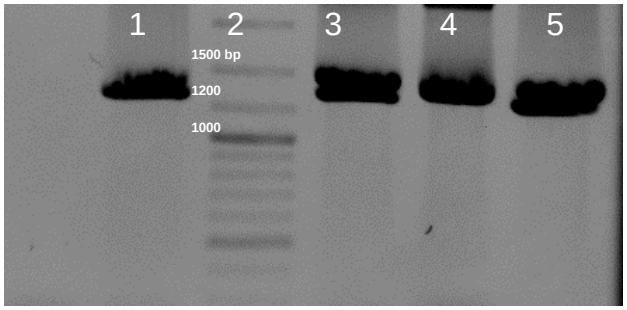


Figure 4-12. Analytical agarose gel of integration of the *baej* fragment in to the pCC1FOS vector containing siphonazole gene cluster.

Agarose gel showing the PCR result using primers for the amplification of a *baeJ* fragment (1243 bp) after transformation *E.coli* BW25113 containing pCC1FOS-siphonazole with *baej*-Zeo fragment. 1: Positive control, plasmid DNA (pGF27+*baeJ*); 2: DNA ladder; 3-5: independent colonies.

After proving the presence of the *baej* fragment in the isolated constructs, the pCC1FOS-siphonazole vector harboring the *baeJ*-Zeo fragment (pCC1FOS-siphonazole+*baej*) should be transferred into *B. amyloliquefaciens* FZb42 according to the protocol 3.11.6. However, the transfer of pCC1FOS-siphonazole+*baej* into *B. amyloliquefaciens* FZb42 was not successful, even when applying different conditions for electroporation e.g. increasing the pulse length, increasing the DNA concentration and using different electroporation buffers.

4.2.1.2 Heterologous expression of the siphonazole gene cluster using a Bacterial artificial chromosome (BAC) based construct

While heterologous expression of small biosynthetic gene clusters (≤ 40 kb) has been used successfully with standard techniques like λ-Red recombination, the approach is not straightforward with large biosynthetic gene clusters (≥ 70 kb) (Heide, 2009; Huang et al., 2012). This might be an explanation why the transfer of the siphonazole biosynthetic gene cluster with a size of approx. 50 kb was unfruitful. Thus, the next step was to change the methodology.

One of the most reliable procedure for transfer of large secondary metabolite gene clusters is using bacterial artificial chromosomes (BACs). These vectors can carry inserts with a length up to 300 kb (Baker & Cotten, 1997; Shizuya et al., 1992). Thus the construction of a BAC vector for cloning and heterologous expression of the siphonazole biosynthetic gene cluster was programed.

pESAC13, *E. coli*-Streptomyces Artificial Chromosomes, the vector which was developed by M. Sosio and S. Donadio, NAICONS, Milano, Italy was used in this study. pESAC13 is a derivative of pPAC-S1, one of the non-conjugative ESACs originally developed (Sosio, Bossi, & Donadio, 2001). The vector contains the *ori*T from the RK2 replicon as well as the ΦC31 *attP*–int system which allows the integration at the chromosomal *attB* site (Kuhstoss & Rao, 1991; Sosio et al., 2000).

The pESAC13 based constructs can be transferred from *E. coli* to a number of *Streptomyces* species as well as into several non-*streptomyces* actinomycetes by conjugation.

Cloning the siphonazole gene cluster to the single copy BAC vector was programed by applying either ligation or λ-Red recombination.

4.2.1.3 Ligation of the siphonazole gene cluster in to pESAC13, First approach

Two *Bam*HI restriction sites flank the siphonazole gene cluster at both sides (upstream of *sphA* and downstream of *sphj*). Thus, it was planned to ligate the complete *siph* cluster into pESAC13 using the *Bam*HI restriction site. Therefore, the pCC1FOS-siphonazole vector and isolated pESAC13 were restricted using *Bam*HI. Restriction mixtures were loaded onto a 0.7% agarose gel and the desired bands were isolated from the agarose gel after electrophoresis (3.12.10). To avoid self-ligation, the linearized pESAC13 vector was subjected to dephosphorylation according to protocol 3.12.7. Afterwards, the linearized pESAC13 was isolated from the dephosphorylation mixture using the DNA clean and concentrator kit.

In the next step the DNA fragment containing the siphonazole gene cluster should be ligated into the pESAC13 (*Bam*HI) vector. The ligation mixture was used to electroporate electro-competent *E.coli* EPI cells (3.11.5). The recovered cells where spread on LB plates supplied with 60 µg/ml kanamycin and 100 µg/ml apramycin for selection.

Despite of applying different strategies like using various insert: vector ratios, different incubation temperatures and different incubation periods no clones could be obtained.

4.2.1.4 Construction of the siphonazole gene cluster in pESAC13, Second approach

Since applying the ligation technique for construction of the siphonazole gene cluster in pESAC13 was unfruitful, applying the λ-Red recombination was planned. This should be done by transformation of *E.coli* cells containing the complete siphonazole gene cluster by the BAC vector as well as two so-called adapters, as it described in the following.

Thus, *E. coli* BW25113 containing pCC1FOS-siphonazole was used as host for the transfer of the linearized pESAC13 as well as the PCR yielded fragments using the λ-Red recombination. These PCR fragments (Adapters) should contain the beginning or the end part of the siphonazole gene cluster as well as homology arms, i.e. sequences showing homology to the pESAC13 backbone.

Therefore, the first primer pair (Start Up- hom and Start da) was designed to contain 30 bp of the pESAC13 sequence at the 5' part, and 20 bp of the beginning part of the siphonazole gene cluster. The second primer pair (End down hom and End up) was designed as mentioned above to contain the 20 bp of the terminal part of the siphonazole gene cluster flanked by 30 bp homologous to the pESAC13 sequence.

Isolated pCC1FOS-siphonazole was used as template for the amplification of the desired fragments. Using the primer pairs (Start Up- hom and Start da) and (End down hom and End up) in the PCR reactions resulted in amplification of adapter1 (962 bp) and adapter2 (891 bp), respectively. Figure 4-13 shows the obtained adapters.

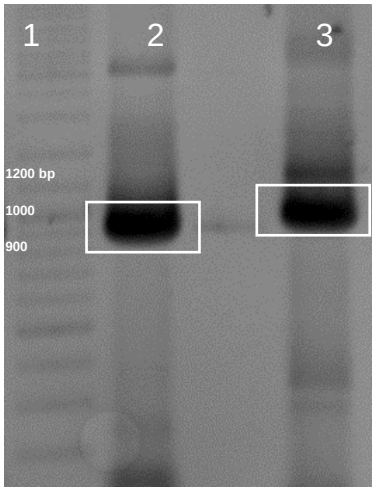


Figure 4-13. Agarose gel with the PCR-derived adapters obtained by amplification with the primer pairs End down hom and End up and Start Up- hom and Start da, respectively.

1: DNA-ladder; 2: Adapter1 carrying 30 bp of pESAC13 and 932 bp of the 5´ start of the siphonazole cluster (with the calculated size of 962 bp); 3: Adapter2 carrying 30 bp of pESAC13 and 951 of the 3´ end of the siphonazole cluster (with the calculated size of 981 bp). The desired bands are highlighted by white boxes.

The PCR-derived adapter1 (962 bp) and adapter2 (981 bp) were isolated from the gel after separation by electrophoresis (Figure 4-13).

pESAC13 (23338 bp) was restricted overnight using the *Bam*HI restriction enzyme. A preparative agarose gel (0.7%) yielded two fragments with the expected size of 20600 bp and 2738 bp, respectively. The bigger band was isolated from the agarose gel for further processing (3.12.10).

Electroporation of *E.coli* BW25113 containing the pCC1FOS-siphonazole vector was done using the PCR-derived adapters as well as the *Bam*HI-restricted vector pESAC13. The electroporation mixture was spread on LB plates containing kanamycin and

ampicillin for selection of positive clones. However, also this approach failed, since no colonies could be obtained on the selective plates.

Also using different concentrations of DNA in the electroporation mixture and applying various incubation temperatures and periods after electroporation did not result in positive clones.

4.2.1.5 Construction of the siphonazole gene cluster in pESAC13, Third strategy

Since all approaches applied so far were unsuccessful, it was planned to assemble the cluster in pESAC13 using λ-Red recombination. In a previous study (Höver. PhD thesis) recombination of three overlapping parts of the siphonazole gene cluster to a single construct was performed by applying λ-Red recombination. Thus, it was planned to transfer three overlapping DNA fragments, containing the start, middle and end part of the siphonazole gene cluster to *E. coli* BW25113 harboring pESAC13.

In the first step *E. coli* BW25113 harboring the pIJ790 helper plasmid was transferred with pESAC13 vector by electroporation. The pIJ790 helper plasmid contains the λ-Red (*gam*, *bet*, *exo*) recombination factors and encodes for a chloramphenicol resistance gene (Gust et al. 2004).

Three different constructs harboring, i) the beginning part of the gene cluster flanked by the streptomycin resistance gene (EC10), ii) the middle part of the gene cluster (EC9), and iii) the terminal part together with an apramycin resistance cassette (IC2) isolated from the corresponding hosts. The isolated fosmids were subsequently restricted using various restriction enzymes.

EC9 was restricted using *Sac*I, EC10 was restricted by *Pci*I and *Pac*I, and *Xma*I and *Dra*III were used for restriction of IC2. Figure 4-14 shows the restriction pattern of EC10, EC9 and IC2 on an agarose gel.

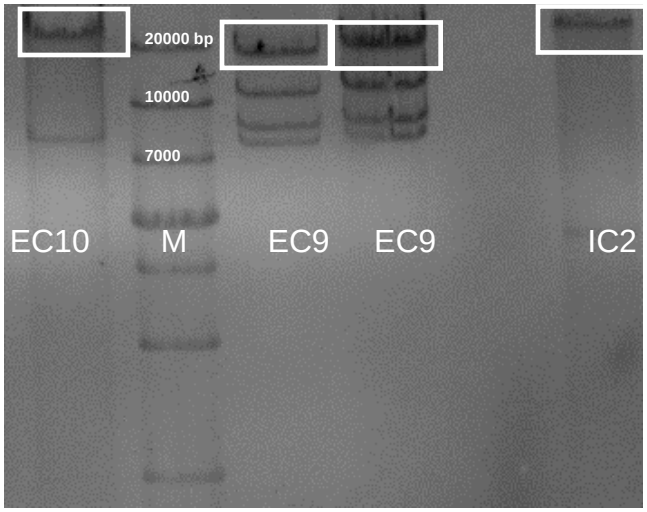


Figure 4-14. Analytical agarose gel of restricted fosmids.

Restriction pattern of the fosmids EC10 (containing the 5´ part of the siphonazole cluster restricted with *Pci*I and *Pac*I), EC9 (containing the middle part of the siphonazole cluster restricted with SacI) and IC2 (containing the 3´ part of the siphonazole cluster restricted with *Xma*I and *Dra*III). The desired fragments carrying overlapping parts of the siphonazole cluster are highlighted by white boxes.

The desired bands were excised from the gel and purified for further processing (3.12.10).

For the assembly of the gene cluster in pESAC13 by homologous recombination two so-called adapters, containing homologous parts of pESAC13 as well as homologous parts of EC10 or IC2, were required. Therefore, two pairs of primers (Start1 BAC and Start2 BAC) and (End1 BAC and End2 BAC) were designed for the amplification of these adapters (Figure 4-15).

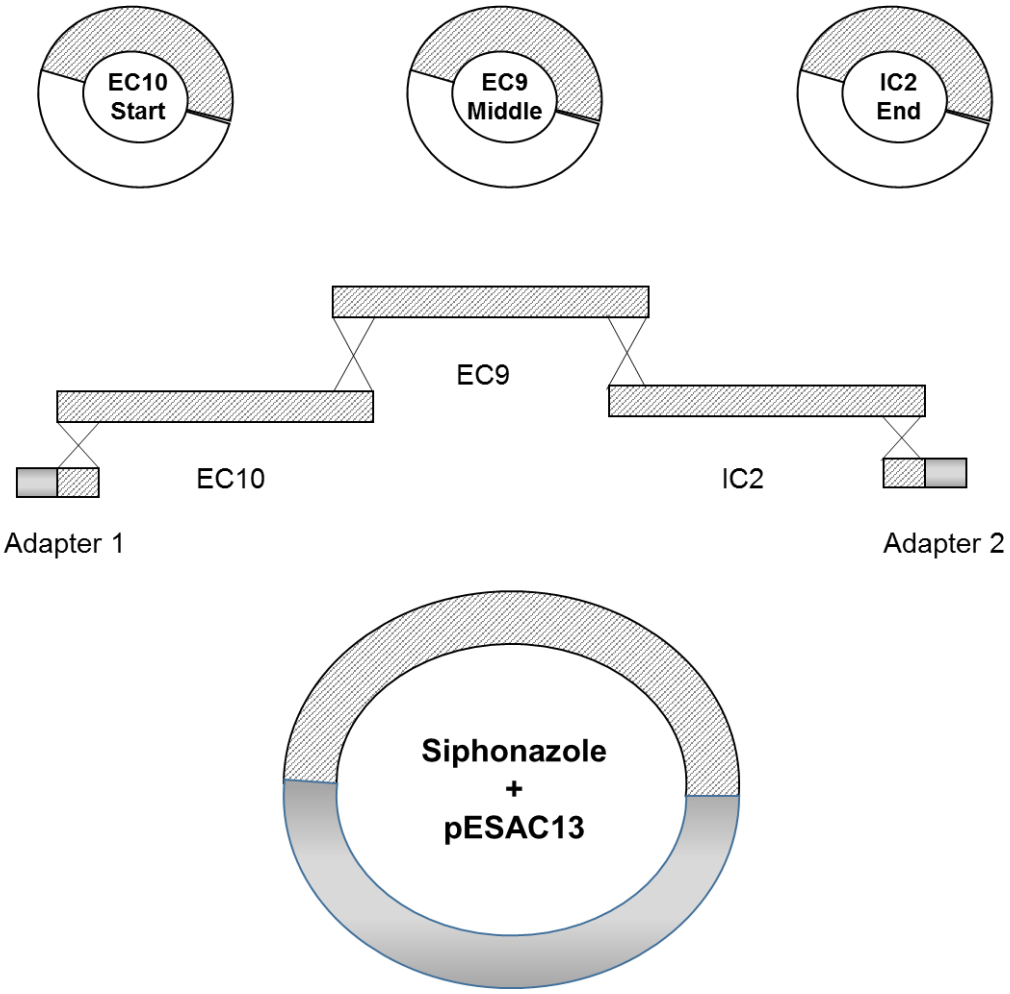


Figure 4-15. Reconstitution of the putative siphonazole cluster in pESAC13 vector using λ-Red mediated recombination.

Linear fragments from EC10, EC9 and IC2 containing the start, middle and end part of siphonazole gene cluster respectively were obtained using the restriction digestion. Two PCR-derived adapters containg the homology sequence of EC10 or IC2 as well as homology arms for the pESAC13 sequence were amplified using the primer pairs Start1 BAC and Start2 BAC and End1 BAC and End2 BAC. Afterward the *E.coli* BW25113 containing the pESAC13 was transferred by the adapters and three linear fragments.

The primer pair Start1 BAC and Start2 BAC was designed for the amplification of a 500 bp fragment showing homology to pESAC13 as well as to EC10 (upstream of the streptomycin resistance gene). The second primer pair End1 BAC and End2 BAC was designed accordingly, containing homology arms for the pESAC13 sequence and for IC2 (downstream of the apramycin resistance gene). In the next step *E. coli* BW25113 harboring pESAC13 was transformed with the three DNA fragments carrying the complete siphonazole gene cluster together with the PCR-derived adapters.

However, no positive colony containing the complete siphonazole gene cluster inserted into the BAC vector, was obtained. Since construction of the siphonazole gene cluster based on the pESAC13 vector was not fruitful either by ligation or applying the λ-Red recombination, it seems that modification of the pESAC13 vector is a most critical step in the heterologous expression of siphonazole.

4.2.2 Heterologous protein expression of domains of the siphonazole cluster

Another approach towards the detailed characterization of siphonazole biosynthesis is the analysis of specific biosynthetic enzyme (domain)s in *in vitro* assays. Of particular interest is the termination of siphonazole biosynthesis, since in the release an unusual decarboxylation and dehydration has to take place. Additionally, the putative DAHP synthase SphI should be tested, to verify the assumption that it is involved in precursor formation. Therefore, these proteins, i.e. the putative hydrolase domain of SphH and the thioesterase SphJ, putatively involved in the unusual decarboxylation and dehydration, as well as the DAHP synthase SphI were heterologously expressed in *E. coli.* The purified enzymes were tested in *in vitro* assays to investigate their exact roles.

4.2.2.1 *In vitro* assays for the investigation of the hydrolase domain (part of SphH) and of the thioesterase SphJ

SphH contains four domains including a putative dehydratase, a ketoreductase, an acyl carrier protein and a hydrolase. Due to the lack of conserved catalytic residues, it is assumed that the dehydratase domain is inactive. Therefore, it can be assumed that the putative hydrolase, or the thioesterase, or both are responsible for the dehydration and decarboxylation in the last step of siphonazole biosynthesis.

To test this hypothesis, the heterologous expression of i) the C terminal part of SphH, consisting of a didomain containing the hydrolase and ACP domain (SphH Hyd-ACP), as well as ii) the thioesterase SphJ was programmed. Afterwards, the catalytic activities of these enzymes should be analyzed using established enzyme assays.

4.2.2.2 The putative hydrolase domain encoded within SphH

Comparative sequence analyses classified the last domain of SphH as a member of the α/β-hydrolase 6 family (pfam12697) (Figure 4-16).

The α/β hydrolase fold is common to several hydrolytic enzymes of widely varied phylogenetic origin and catalytic functions, including thioesterase, peptidase, general hydrolase, and lyase activities. These enzymes can be characterized by a central eight-stranded β-sheet surrounded by α-helices. Hydrolases of this superfamily typically share only low levels of sequence identity and encompass diverging catalytic activities (Ollis et al., 1992; Auldridge et al., 2012).

Despite of various catalytic functions within this family, they often encompass crucial active-site residues forming a Ser-His-Asp catalytic triad for general acid-base and nucleophilic catalysis and an adjacent oxyanion hole for transition state stabilization (Nardini & Dijkstra, 1999; Auldridge et al., 2012).

Unexpectedly SphH-Hyd lacks one of the key active-site residues presumed to be essential for hydrolase activity. Sequence comparisons indicate that the catalytic triad of SphH-Hyd consists of Cys-His-Asp, instead of the traditional α/β-hydrolase triad Ser-His-Asp (Figure 4-16).

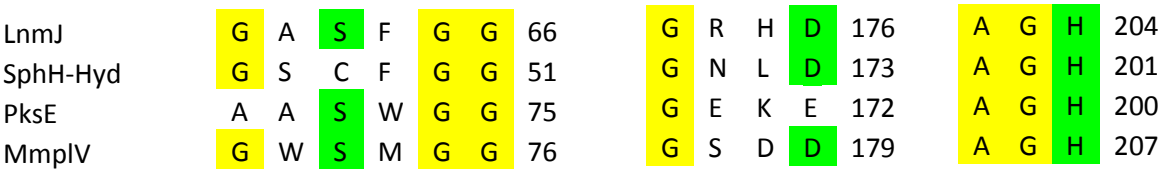


Figure 4-16. Multiple sequence alignment of the hydrolase part of SphH. The crucial active site residues forming a Ser-His-Asp catalytic triad are shown in green. Residues corresponding to the consensus motifs are coloured yellow. It can be seen that SphH-Hyd carries a Cys residue instead of the conserved Ser residue. The reference sequences are: LnmJ from *Streptomyces atroolivaceus* (GI: 26541536); PksE from *Bacillus subtilis* (GI: 111052871) and MmplV from *Cellvibrio japonicus* Ueda107 (GI: 190688108).

4.2.2.3 Heterologous expression and purification of the didomain SphH Hyd-ACP

For the expression of the didomain construct SphH Hyd-ACP (3.12.13), *E. coli* BAP1 cells harboring pET151-SphH were used. The desired protein was overexpressed and purified according to the method 3.13.3 from 1 liter culture. The purification of SphH Hyd-ACP as a His-tagged protein was successful. A prominent band corresponding to the expected size of 53.5 kDa was enriched during purification. This was analyzed using SDS-PAGE analysis (Figure 4-17).

The resulting protein was subsequently concentrated and re-buffered in Tris HCl buffer (pH7).

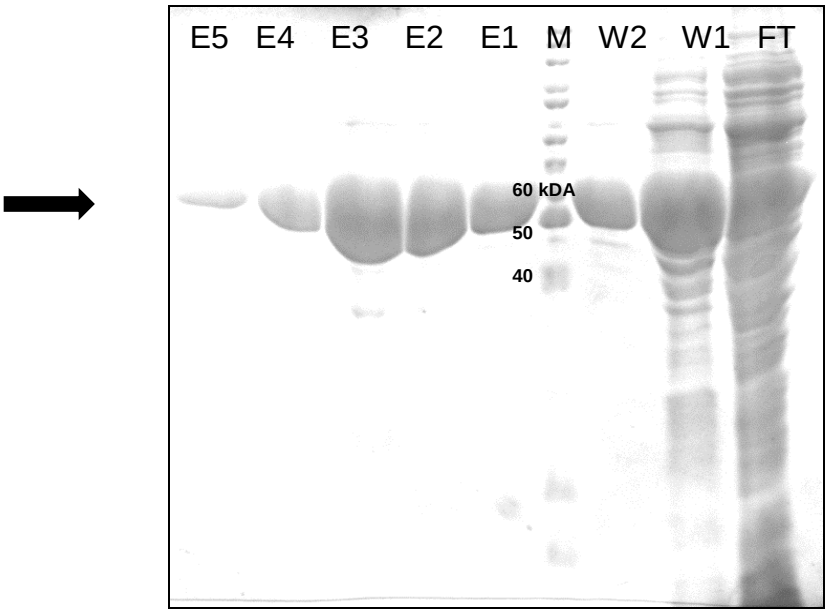


Figure 4-17. SDS-PAGE of a typical purification of SphH Hyd-ACP by affinity-chromatography on a Ni-NTA column.

FT, flow through; W1 and W2, (washing steps with 30 and 50 mM imidazole, respectively); E1–E5, elution fractions (100-350 mM imidazole) once at 100 mM, once at 150 mM, once at 200 mM, and twice at 300 mM).; M, marker. Black arrow indicates calculated mass of 53.5 kDa for the SphH Hyd-ACP.

4.2.2.4 The putative thioesterase, SphJ

Blast search annotated SphJ as a thioesterase (pfam00975). Protein members of this family have been shown to have similarity to the type II fatty acid thioesterases of vertebrates. TE domains are responsible for chain termination and release has been well characterized from PKS, NRPS, and PKS/NRPS systems (Kohli et al., 2002; Tang et al., 2004). Similar to members of the α/β-hydrolase 6 family (like SphH-Hyd) they contain an α/β-hydrolase fold domain and also share the same G-x-S-x-G motif around a catalytic serine (Eys et al., 2008). The members of this family possess the active-site

Ser and His residues as part of the highly conserved G-X-S-X-G and G-X-H-F motifs, respectively.

A multiple sequence alignment was performed with similar TE domains to identify the signature sequences *in silico*. SphJ encompasses the highly conserved G-X-S-X-G motif (Figure 4-18), while in the second motif phenylalanine was replaced by methionine in all of the TE example sequences.

B.licheniformis	G	H	S	M	G	G	92	G	G	H	M	220
B.subtilis	G	H	S	M	G	G	89	G	G	H	M	217
B.altitudinis	G	H	S	M	G	G	90	G	G	H	M	218
SphJ	G	H	S	L	G	G	90	G	P	H	M	218
B.cereus	G	H	S	M	G	G	87	G	G	H	M	215

Figure 4-18. Multiple sequence alignment of the thioesterase SphJ.

Highly conserved G-X-S-X-G and G-X-H-F motives are shaded in yellow. Active-site Ser and His residues are shown in green. The reference sequences were taken from thioesterase of *Bacillus licheniformis* (GI: 3080745); the same protein of *Bacillus altitudinis* (GI: 728915111) as well as *Bacillus subtilis* (GI: 668753059), and lichenysin synthetase D of *Bacillus cereus* 172560W (GI: 228601438).

4.2.2.5 Heterologous expression and purification of SphJ

The expression and purification of the recombinant protein SphJ was accomplished in an analogous way to the expression of SphH Hyd-ACP.

The His-tagged SphJ was purified using affinity chromatography. Analyzing the different fractions of the purification process on a SDS-PAGE revealed a band in the elusion fractions which corresponds to the expected size of SphJ (31.2 kDa) (Figure 7-19).

The protein concentration, determined by UV spectroscopy (Thermo Scientific nanodrop), of a typical purification was 7 mg/ml.

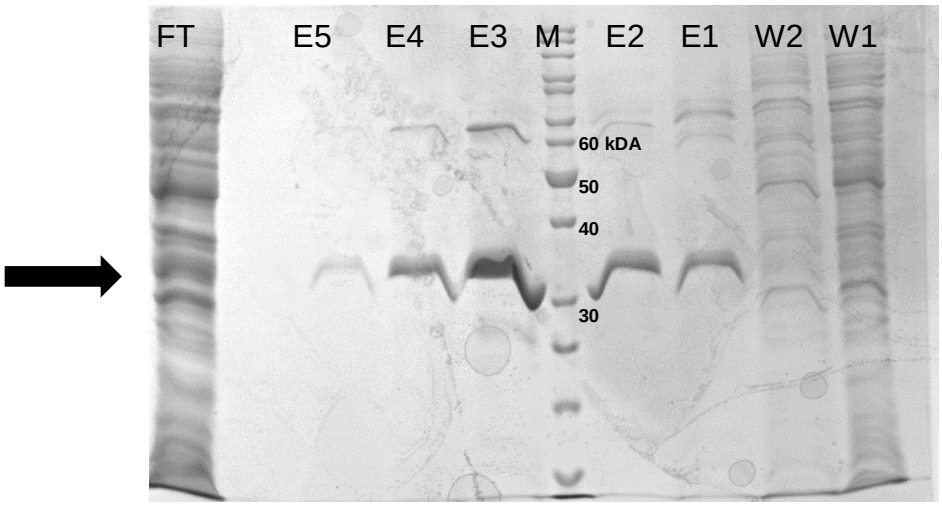


Figure 4-19. SDS-PAGE of a typical purification of SphJ by affinity-chromatography on a Ni-NTA column.

FT, flow through; W1 and W2, (washing steps with 30 and 50 mM imidazole, respectively); E1–E5, elution fractions (100-350 mM imidazole); M, marker. The black arrow indicates the calculated mass of 31.2 kDa for the SphJ.

4.2.2.6 *In vitro* assays to prove the functional role of Hydrolase domain SphH and SphJ

To investigate the unusual dehydration and decarboxylation as the final step in siphonazole biosynthesis several *in vitro* experiments were performed.

There are two possibilities for analyzing the conditions necessary for these unusual reactions.

The first one is analyzing the molecule while it is attached on the 4′-phosphopantetheine arm of the ACP. This can be done by using multi-stage tandem MS in the so-called phosphopantetheine (PPant) ejection assay. Thus, the molecular mass of substrates loaded on to the carrier proteins can be measured. Therefore the catalytical activity of the associated enzyme can be deduced from the mass of the corresponding PPant fragments (Meluzzi et al., 2008).

In the second approach analysis of the off-loaded molecule can be envisaged. This off-loading can be achieved by two possibilities i) by chemically hydrolyzing the thio-ester while the substrate is loaded to the PPant arm, by using 0.1 M KOH, or ii) by applying the heterologously expressed TE domain. TE domains mostly play the role of the termination domain in natural product biosynthetic clusters, by catalyzing the breakage of the thioester linkage between the final product and the PPant arm of the ACP after the final elongation step. In the second assay hetrologously expressed TE was either used separately or in combination with the hydrolase. In that way, this experiment should shed light on the possibility if the TE domain itself is involved in dehydration and decarboxylation, or if both enzymes are required for generating the final product.

To perform the assays, a suitable substrate which can be loaded to the (phosphopantetheinylated) ACP was needed. The substrate should comprise two features first of all it should be an analogue to the respective intermediate of siphonazole biosynthesis, and it also should be able to be loaded this substrate onto the phosphopantetheinylated arm of the expressed *holo*-ACP domain. Thus, a substrate mimicking the respective intermediate of siphonazole biosynthesis was synthesized, i.e. substrate **2** in figure 4-20, as an *N*-acetylcysteamine (SNAC)-coupled molecule.

Further, the active *holo*-ACP form is required in this experiment. Therefore, *E. coli* BAP1 cells were applied as the heterologous expression host, since this genetically modified strain carries the *sfp* gene from the surfactin biosynthetic gene cluster of *Bacillus subtilis*, which encodes a promiscuous phosphopantetheinyl (PPant) transferase. This enzyme catalyzes the posttranslationally modification of the ACP by transferring of the PPant arm to the conserved serine residue of the *apo*-ACP (Quadri et al., 1998).

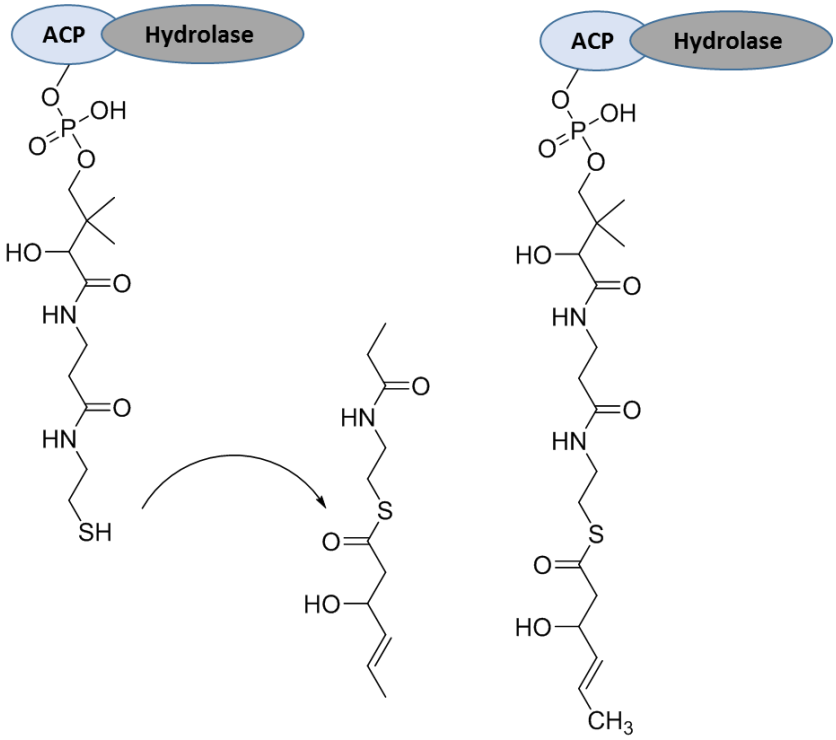
4.2.2.7 Phosphopantetheine (PPant) ejection assay

To deduce the catalytical function of the putative hydrolase domain of SphH the PPant ejection assay was performed. The principle of the assay is analysis of carrier bound intermediates in a MS based *in vitro* assay (Meluzzi et al., 2008).

In the first step the SNAC-activated substrate should be attached to the PPant residue of the *holo*-ACP (Figure 4-19). In the second step the catalytical activity of the putative SphH can be analysis by the PPant ejection assay while the substrate is loaded to PPant arm of the ACP.

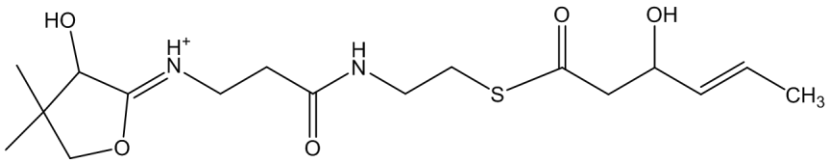
Thus, In brief, freshly purified SphH Hyd-ACP was incubated with SNAC-activated substrate (Figure 4-20). The coupling reaction was done by Sarah Frank (See 3.16.1). The later **2** represents a simplified mimic of the natural substrate. After incubation, the reaction was stopped by injection of the complete assay mixture into an HPLC device (See 3.16.2). The purified protein was collected manually and dried. To analyze the assay, the protein sample was subjected to mass spectrometry.

Collision energy in the linear ion trap applied to such a *holo*-protein resulted in the ejection of the PPant residue. If a substrate is tethered to this PPant arm, ejection ions will consist of PPant+substrate. Two possible ejection ions which can could be expected in this experiment are shown in figure 4-21.



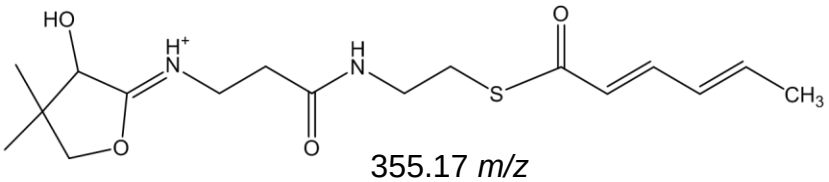
2 **3**

Figure 4-20. Scheme for the binding of substrate **2** onto the phosphopantetheine (PPant) arm of SphH Hyd-ACP.



373.18 *m/z*

4



355.17 *m/z*

5

Figure 4-21. Structure of expected ejection ions of SphH.

When the PPant arm plus attached substrate is liberated from the *holo*-protein ACP by applying collision energy, and resulted in an ejection ion. Formation of the diene in ejected ion **5** was expected to be performed in presence of SphH Hyd-ACP while the inactive enzyme can lead to detection of ejected ion **4**.

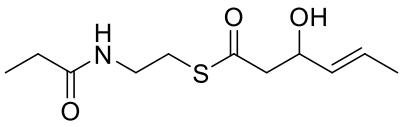
The MS measurement revealed no ejection ions. Therefore, using this approach no insights into the catalytic role of SphH Hyd-ACP could be obtained.

4.2.2.8 Analyses of the ejected molecule by GC-MS

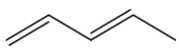
Another possibility for the determination of the decarboxylation and dehydration process would be the analysis of off-loaded reaction products. The off-loading of an ACP-attached molecule can be achieved either by chemically hydrolyzing the thio-ester through which the substrate is attached to the PPant arm, i.e. by using 0.1 M KOH, or by using the catalytic activity of a TE.

Thus, the reaction mixtures containing freshly purified SphJ (the TE) and SphH Hyd-ACP or only TE domain, as well as substrate **2** were set up. The reactions were stopped by the addition of methanol and the reaction products were extracted using ethyl acetate. It was assumed that substrate **6** should be present in the extract (Figure 4-22).

Further, the same assay mixture as described was prepared, but after incubation the reaction was treated with KOH to chemically liberate the ACP-tethered products.



2



6

Figure 4-22. Molecules expected in the assay using SphJ and SphH.

The off-loaded substrate **6** was expected to be construct from substrate **2** using 0.1 M KOH, or by using the catalytic activity of a TE.

As a control two mixtures containing the authentic standard (*trans*-1,3-pentadiene) for the expected product with or without protein were prepared. These mixtures were incubated and extracted using the same conditions as for the reaction mixtures.

For analysis the extracted products were directly injected to the gas chromatography–mass spectrometry instrument.

Injection of the extracted standard from the reaction mixture without the protein solution resulted in the peaks at 8.43 and 7.75 min with the mass of 68.062 D as expected for compound **6**. The mentioned peaks were proven the precence of the authentic standard either in (E)-penta-1,3-diene or (Z)-penta-1,3-diene form, although determination of the right form for the detected peaks could not be judge from this experiment. However, these peaks were not observed in the reaction mixture containing the protein (Figure 4-23).

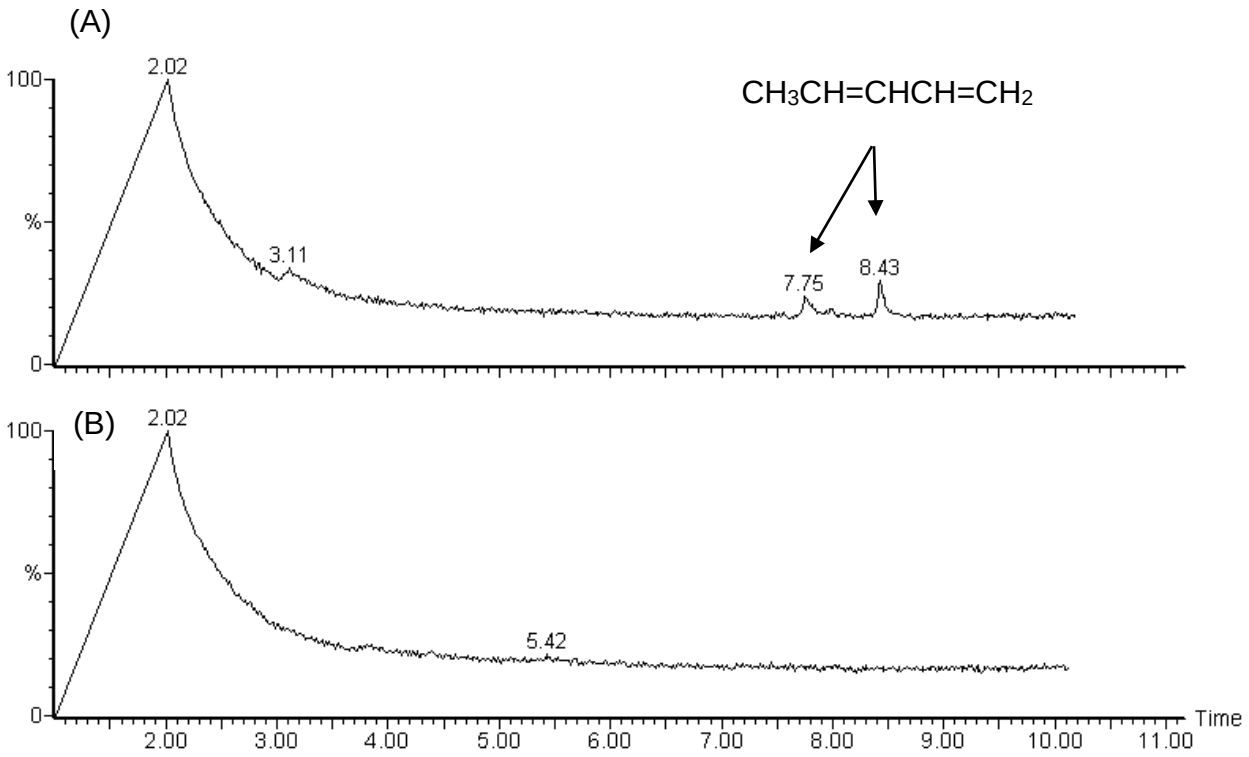


Figure 4-23. GC chromatogram of (A): the extracted standard from the reaction mixture containing the standard and buffer without the protein solution and (B): extracted standard from the reaction mixture containing the standard, buffer and protein solution.

Efforts for detection of the released products, either in the enzymatic process in presence of both proteins as well as the single TE domain or in chemically hydrolyzing process using KOH, failed (Appendix 7.3).

4.2.3 Starter Unit – DAHP Synthase SphI

4.2.3.1 The putative aldolase domain encoded by SphI

The alone-standing gene coding for SphI is located directly downstream of *sphH* and upstream of *sphJ*.

Conserved domain database annotated SphI as DAHP synthase I family (EC: 2.5.1.54). DAHPS, generates DAHP and inorganic phosphate from the condensation of PEP and E4P (Wu et al., 2003). Three conspicuous motifs which may correspond to active-site residues of DAHP synthase, G-P-C-S, K-P-R-T-S/T, and I-G-A-R are shown in figure 4-24. Missense mutants within the G-P-C-S and I-G-A-R motifs of *E. coli* AroAW were reported, whereby either loss of catalytic activity or loss of feedback resistance were found interspersed (Ray et al., 1988).

Database search indicates K-P-R-T as the catalytic motif of the DAHPS members of the family I, while eight family II (plant type) DAHPS possess the K-P-R-S motif in the same region (Subramaniam et al., 1998).

S.thermophilus	G	P	C	S	104	K	P	R	T	136	I	G	A	R	187
N.hollandic	G	P	C	S	104	K	P	R	T	136	I	G	A	R	187
C.bacterium	G	P	C	S	104	K	P	R	T	136	I	G	A	R	187
Anaeromyxobacter sp	G	P	C	S	103	K	P	R	T	135	I	G	A	R	186
H.contractile	G	P	C	A	105	K	P	R	T	137	I	G	A	R	188
SphI	G	P	C	S	104	K	P	R	T	136	I	G	A	R	187

Figure 4-24. Multiple sequence alignment of the DAHP synthase SphI. Three conspicuous motifs G-P-C-S, K-P-R-T-S/T and I-G-A-R which are proposed to be important for catalytic activity of DAHPS are shaded in yellow. The reference sequences are taken from *Sphaerobacter thermophilus* DSM 20745 (GI: 269787160), *Nitrolancea hollandica* Lb (GI: 390172899), *Clostridiales bacterium* DRI-13 (GI: 736402457), *Anaeromyxobacter* sp. PSR-1 (GI: 775301777), and *Haloplasma contractile* SSD-17B (GI: 543105077) (Subramaniam et al., 1998).

Crystal structures suggest that the active site metal binds to the four residues in *A. aeolicus* KDO8P synthase (Cys^{11}, His^{185}, Glu^{222}, and Asp^{233} ligands to Cd^{2+}) and *T. maritima* DAHP synthase (Cys^{102}, His^{272}, Glu^{298}, and Asp^{310}). The same active site residues are absolutely conserved in SphI (Cys^{103}, His^{273}, Glu^{299} and Asp^{310}) (Figure 4-25) (Wu et al., 2003; Duewel et al., 2001; Igor A. Shumilin et al., 2002).

S.thermophilus	G	P	C	S	104	S	H	G	T	275
N.hollandic	G	P	C	S	104	S	H	G	T	275
C.bacterium	G	P	C	S	104	S	H	G	T	275
Anaeromyxobacter	G	P	C	S	103	S	H	G	I	274
H.contractile	G	P	C	A	105	S	H	A	T	276
T. maritima	G	P	C	S	103	S	H	S	G	274
SphI	G	P	C	S	104	S	H	G	V	275

S.thermophilus	S	D	G	311	V	E	V	300
N.hollandic	S	D	G	311	L	E	V	300
C.bacterium	S	D	G	311	V	E	V	300
Anaeromyxobacter	S	D	G	310	V	E	V	299
H.contractile	S	D	G	312	I	E	V	301
T. maritima	S	D	G	310	V	E	V	299
SphI	S	D	G	311	I	E	V	300

Figure 4-25. Multiple sequence alignment of the DAHP-synthase SphI (metal binding residues).

Putative metal binding residues were shaded (Cys^{102}, His^{272}, Glu^{298}, and Asp^{310}) based on amino acid sequence for DAHP synthase of *T. maritima* in green. The reference sequences are taken from *Sphaerobacter thermophilus* DSM 20745 (GI: 269787160), *Nitrolancea hollandica* Lb (GI: 390172899), *Clostridiales bacterium* DRI-13 (GI 736397190*), Anaeromyxobacter* sp. PSR-1 (GI: 775301777), *Haloplasma contractile* SSD-17B (GI: 543105077), And *Thermotoga maritima* MSB8 (GI: 37999819) (Wu et al., 2003).

4.2.3.2 Heterologous expression and purification of the DAHP synthase SphI

In the first attempt the corresponding strain containing the (pET151-SphI) construct, which obtained from the previous work by T. Höver, was inoculated in 1 liter LB medium supplied by ampycilin for selection. The bacterial culture was incubated at 37°C and after cooling down was induced by IPTG (0.5 mM). In this way, most of the purified protein was located in the insoluble fraction of the pellet (Figure 4-26 A). To overcome this, a different incubation temperature as well as lower IPTG concentration were tested. Thus, the bacterial culture was incubated at 16°C and induced by of IPTG (0.4 mM) which resulted in higher concentration of the soluble protein (Figure 4-26 B). The expressed His-tagged protein (41.3 kDa) was purified by affinity chromatography on a NI-NTA column using the slightly modified protocol 3.13.3 (Using higher concentration of imidazole in washing and elution steps).

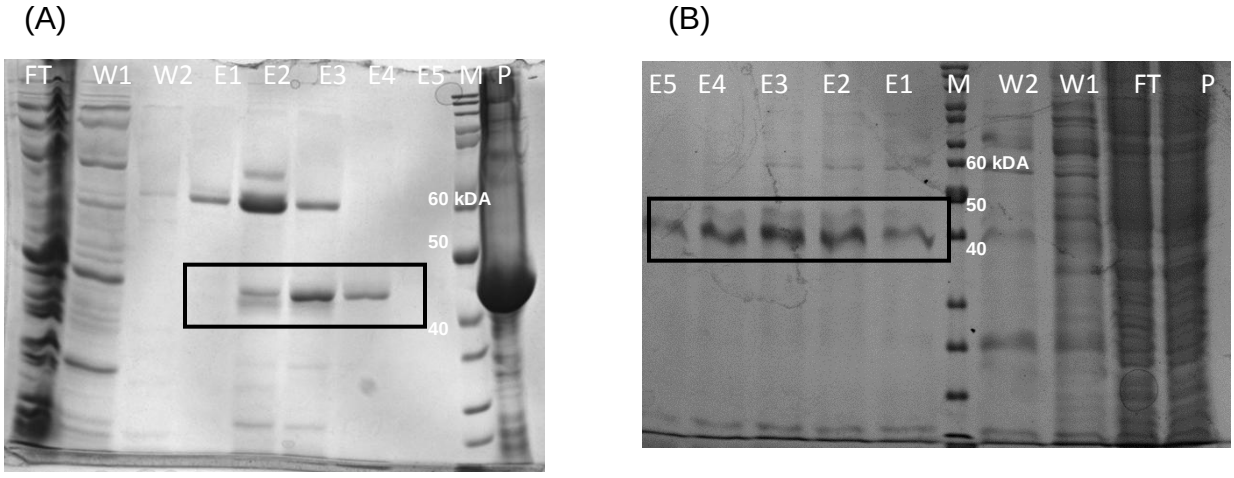


Figure 4-26. SDS-PAGES of a typical SphI purification by affinity-chromatography on a Ni-NTA column.

(A) FT: flow through; W1 and W2, (washing steps with 30 and 50 mM imidazole, respectively); E1–E5: elution fractions (100-300 mM imidazole); M, marker; P: pellet. (B) FT: flow through; W1 and W2, (washing steps with 50 and 100 mM imidazole,

respectively); E1–E5, elution fractions (150-300 mM imidazole); M, marker; P: pellet. Black boxes indicate the calculated mass of 41.3 kDa for the SphI.

The purified enzyme was always used freshly in the reaction mixture. The enzyme was proved to be active either by detection of the dropping absorption at 223 nm which reflects the PEP consumption in a continuous spectrophotometric method (3.15) or by observing the production of DAHP using the LC-ESI MS (3.14.3). Applying these assays for the reaction standard mixtures containing no protein resulted in no dropping UV absorption, as well as no DAHP production.

Hence, further experiments for the determination of the optimal conditions for SphI activity were performed, using different temperature and pH values. In addition the metal requirement, as well as the kinetic constants of SphI were analyzed.

4.2.3.3 Effect of temperature on the enzymatic activity of SphI

Determination of the enzyme activity was measured between 10°C and 70°C by applying an assay according to method 3.15.2. The reaction mixture was incubated at least for 5 minutes at the desired temperature and the reaction was started by adding freshly purified enzyme. The enzyme activity was monitored spectrophotometrically at 232 nm.

Measurement were performed in duplicate. 100% activity are equivalent to a consumption of 80 µM PEP in 30 minutes. The temperature optimum for SphI was shown to be 30°C (50% and 100%). Half of the maximum activity (about 20%) was observed at 23°C (50% and 25%). The enzyme was still slightly active at 40°C (0% and 25%), but no activity was observed at 10°C, and also no activity remained at 50-70°C.

4.2.3.4 pH dependence of the enzymatic activity of SphI

pH dependence of the enzymatic activity was determined by measuring the activity between pH 5-9. Enzyme activity was measured at room temperature using BTP buffer (pH 5-9) applying the spectrophotometric assay for 30 minutes.

pH optimum for SphI was observed at pH 7. At pH 6 and 8 activity was measured to be around 12.5%. Figure 4-27 shows that at a pH of 5, and 9, respectively, no activity of the DAHP synthase was observed.

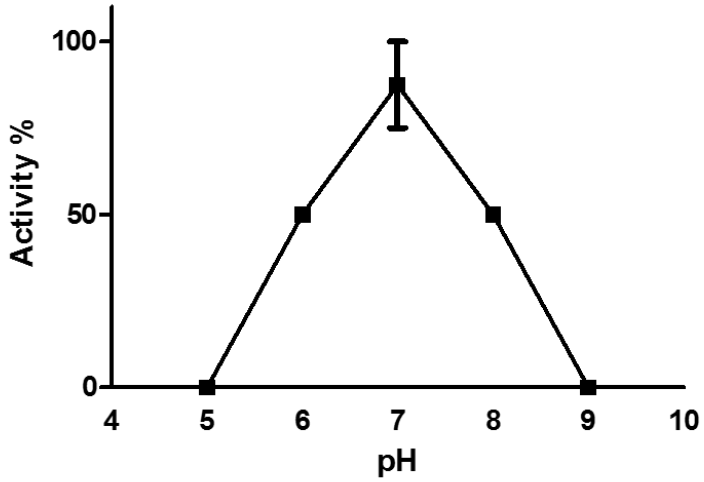


Figure 4-27. pH dependence of the activity of SphI.

Disappearance of PEP in the reaction mixture containing 80 µM PEP, 350 µM E4P, 100 µM $MnSO_4$ and BTP buffer (pH 5-9) were observed by adding SphI, at room temperature and the values were recorded after 30 minutes from the beginning of reaction (Method 3.15). 100% activity are equivalent to a consumption of 80 µM PEP in 30 minutes. Error bars correspond to the standard deviation of two determinations. The mean value of two independent protein purifications is given (pH 5: 0±0%, pH 6: 50±0%, pH 7: 87.5±12.5%, pH 8: 50±0 and pH 9: 0±0%).

4.2.3.5 Metal requirements of SphI

Results deduced from the multiple sequence alignment (Figure 4-25) pointed toward the metal requirements of SphI. Thus the effect of different metal ions (Mn^{2+}, Mg^{2+}, Zn^{2+}, Cu^{2+} and Cd^{2+}) on SphI activity was examined. The highest activity was observed with Mn^{2+}. However, *in vitro* activity of SphI was also observed using other divalent ions. The activity decreased in the following order: $Cu^{2+} > Mg^{2+} > Cd^{2+} > Zn^{2+}$.

The addition of EDTA to the reaction mixture abolished SphI activity completely. Thus, the results of the sequence alignment classifying SphI as a metaloenzyme was verified by the *in vitro* assay (Table 4-4).

Table 4-4. SphI activity under different assay copnditions.

Metal ions	[a]Activity (unit)
$MnCl_2$	0.58±0.08
$MgCl_2$	0.33±0.17
$ZnCl_2$	0.08±0.08
$CuSO_4$	0.41±0.08
$CdSO_4$	0.16±0.00
EDTA	0.0±0.0

[a]One unit of enzyme activity is defined as the consumption of 1 µmol of PEP per minute. To investigate the metal dependence of the DAHP synthase, the enzyme and the assay mixture were separately incubated with EDTA for 10 minutes prior reaction. The reaction was started by adding EDTA treated SphI to the reaction mixture (80 µM PEP, 350 µM E4P, EDTA 100 µM, 50 mM BTP buffer; pH 7). Determination of metal ions on SphI activity was obtained in the reaction mixtures by substituting different divalent cations, i.e. Mg^{2+}, Zn^{2+}, Cu^{2+}, and Cd^{2+}, for Mn^{2+} (100 µM). Assay was performed as described in method 3.15. The mean value of two independent protein purifications

obtained after 30 minutes of incubation of the reaction mixtures with SphI, are given in table 4-4.

4.2.3.6 Kinetic properties of SphI

The obtained optimum conditions (pH 7, 30°C and addition of $MnCl_2$) were used in this study for the identification of kinetic properties of SphI. Using Graf Pad Prism 5, the kinetic constants of SphI were determined according to double-reciprocal curves.

Double-reciprocal (Lineweaver-Burk) plot of SphI was calculated against the substrate concentration. Thus, various concentrations of E4P (0.025-1.000 mM) were added to the reaction mixture containing steady concentration of PEP (80 µM). Michaelis–Menten kinetics contents (V_{max}, K_m and K_{cat}) of the SphI was determined and, is presented in table 4-5.

Table 4-5. Kinetic parameters of SphI achieved by the system, at saturating substrate concentrations. The mean value of two independent protein purifications is given.

V_{max} (U/mg^{-1})	K_m (mM)	K_{cat} (s^{-1})
$0.19 \times 10^{-3} \pm 0.1 \times 10^{-3}$	0.3779±0.0208	0.0088±0.0003

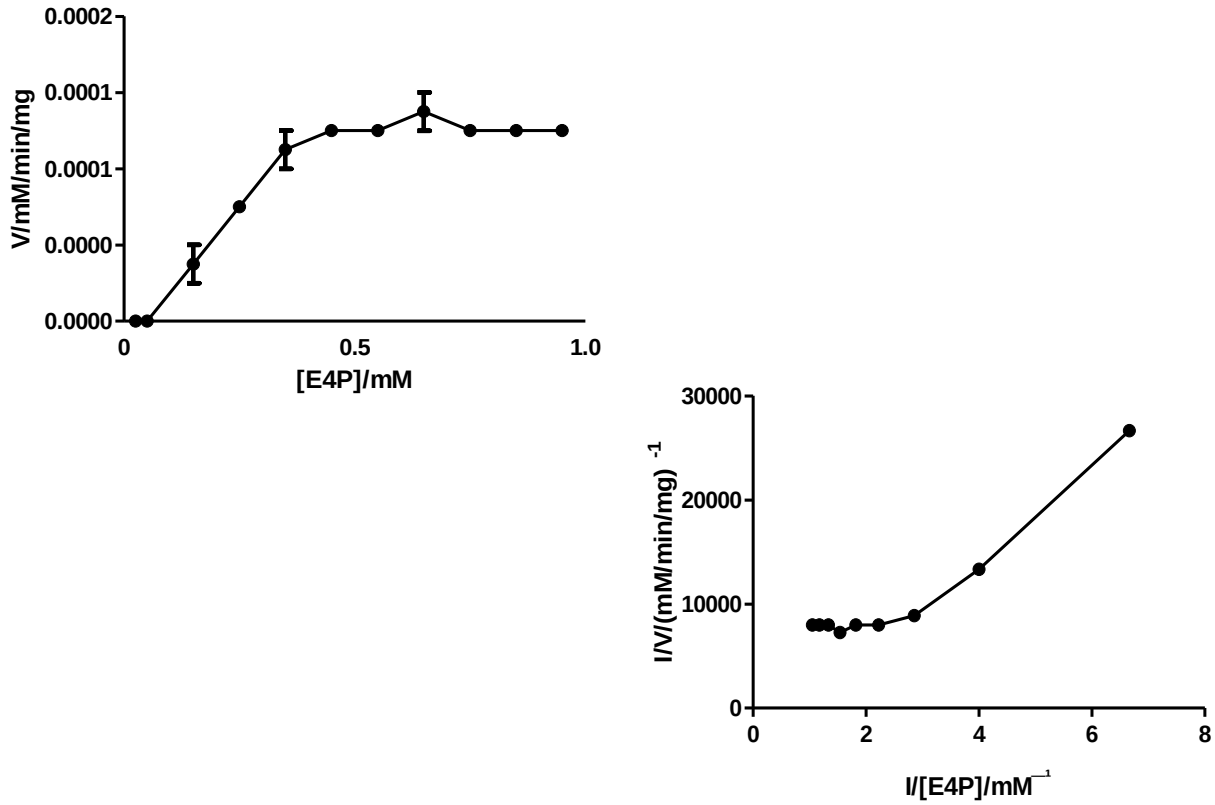


Figure 4-28. Lineweaver-Burk plot of DAHP (SphI) synthase reaction rate against substrate concentration. The erythrose-4-phosphate concentrations was varied (0.025-1.000 mM) while phosphoenolpyruvate concentration maintained at 0.08 mM. The reaction mixtures containing 100 µM $MnSO_4$, were initiated by addition of SphI. Reaction systems were incubated for 30 minutes at room temperature. The UV absorption at 232 nm was recorded after 30 minutes of the adding SphI for each reaction. The mean value of two independent protein purifications is given.

4.2.4 DISCUSSION

4.2.4.1 Heterologous expression of the siphonazole gene cluster using a BAC-based construct

Since *Herpetosiphon* sp. 060 seems to be unamenable to genetic manipulation, heterologous expression of the siphonazole gene cluster could help gaining more insights into some of the still unravelled processes in siphonazole biosynthesis. As the heterologous expression of the reconstituted complete putative siphonazole gene cluster in *E.coli* cells did not result in production of siphonazole (Höver. PhD thesis), other expression hosts were considered as a possible hosts for the heterologous expression (Höver. PhD thesis). Thus, the choice of a suitable host strain was the first critical step in this study.

B. amyloliquefaciens FZb42, due to its similarity of the codon usage with the original host, was the first appropriate candidate for heterologous expression. On the other hand *Streptomyces coelicolor* A3(2) was shown to be a prolific producer of secondary metabolites. Therefore, despite large differences in GC-content between these two bacteria, this strain may also be used in this study (McDaniel et al., 1993). Previous studies showed that differences in GC-content not necessarily are the limiting factor for using this stain as a foreign host, since heterologous expression of the cyanobacterial protein kinase C activator lyngbyatoxin A in *S. coelicolor* was successful, despite to the GC-content difference (Jones et al., 2012).

In the first approach *B. amyloliquefaciens* FZb42 was chosen as the possible expression host. Therefore, integration into the genome was planned using homologous recombination into the gene *baej*. Due to the successful integration of the smaller construct (pGF27+*baeJ*) into the genome of *B. amyloliquefaciens* FZb42 (Figure 4-11), an analogous approach was started for the heterologous expression of the siphonazole gene cluster in *Bacillus*. Therefore, the PCR fragments containing this gene as well as zeocin® as the resistance marker gene were cloned in the pCC1FOS-siphonazole vector, using λ-Red recombination.

Although the same established process was applied for the transformation, several transformations trials failed. This can be caused by instability of the construct. The fosmid vectors are capable to carry an insert up to 40 kb, while the siphonazole gene cluster is about 50 kb. Therefore, the fosmid-based construct (pCC1FOS- Siphonazole) which is 71.275 bp can be instable during the transformation.

A method for overcoming this problem can be the usage of bacterial artificial chromosomes. These vectors can be applied for the stable cloning of large DNA up to 350 kb (Kakirde et al., 2011; Huang et al., 2012). Heterologous expression of secondary metabolites i.e., meridamycin which biosynthetic gene cluster is 90 kb in size, was successful using a BAC vector (H. Liu et al., 2009). Therefore, reconstruction of the gene cluster in a single copy BAC vector (pESAC13) was planned.

Ligation of the siphonazole gene cluster in pESAC13 vector resulted in no colony on the selective agar. This can be due to the large size of the DNA-fragment (~ 50 kbp) which should be cloned into the BAC vector. It was shown that assembling large fragments by restriction and ligation is inefficient and time-consuming. Therefore, λ-Red mediated recombination which was shown to be a reliable method to obtain the BAC-based construct was applied in this study (Gust et al., 2004).

The fosmids EC9, EC10 and IC2 containing the complete siphonazole gene cluster, were identified to have overlapping sequences. This feature was used for reconstructing them based on the BAC vector once again using the λ-Red recombination. Therefore, the restricted BAC vector as well as the adapters containing sequences showing homology to the start and end part of the siphonazole gene cluster as well as 30 bp homology regions to the BAC vector sequence were transferred to the *E.coli* BW25113 harboring pCC1FOS-siphonazole vector. However, none of the applied methods resulted in any colonies.

As all the described methods for manipulation of the BAC vector to get the BAC-based construction failed, the only probable way would be the construction of a new genomic BAC library of *Herpetosiphon* sp.060.

The BAC vectors can carry inserts of up to 350 kbp. Therefore, the chance of obtaining a single BAC-based construct containing the complete siphonazole gene cluster is high, as it was shown in the following examples (Stone et al., 1996; Farrar & Donnison, 2007).

The generation of a genomic library of *Streptomyces tsukubaensis* NRRL 18488 producing FK506, a high value calcineurin inhibiting immunosuppressant, resulted in two clones containing the FK506 complete gene cluster (83.5 kb) (Jones et al., 2013).

In addition, screening of the constructed genomic library of the chaxamycin producer *S. leeuwenhoekii*, using a BAC vector was also successful, yielding a BAC-based construct containing the complete chaxamycin gene cluster (80.2 kb).

pESAC13 was developed to be a shuttle vector between *E. coli*, where it replicates autonomously, and a suitable *Streptomyces* host. The transfer can be done by conjugation. This vector contains the phage P1 origin of replication, as well as the phage phiC31 integration system that allows stable insertion into the genomes of most *Streptomyces* species by site-specific genomic integration (Jones et al., 2012; Thyagarajan et al., 2001). To increase the host spectrum of this vector towards bacilli, an internal gene, i.e. *baej*, encoding the phosphopantetheine binding protein BaeJ, or *amyE*, which encodes α-amylase, should be integrated to the present vector. This modification can readily be performed by the implementation of recombination methods like λ–Red. Afterward, the newly modified vector should be tested and then used in the future for the construction of a genomic library of *Herpetosiphon* sp.060.

4.2.4.2 Decarboxylation and dehydration of the terminal acetate moiety

The feeding experiments with [1-^{13}C]acetate revealed clearly that the last incorporated acetate unit in siphonazole biosynthesis undergoes unusual decarboxylation and dehydration. The formation of a similar terminal alkene moiety was investigated in the biosynthesis of curacin A by the Sherman group (Gu et al., 2009). Curacin A is a marine cyanobacterial metabolite with anticancer activities. It acts by interacting with the colchicine binding site on tubulin and thereby inhibits microtubule polymerization. The

compound was isolated from the marine cyanobacterium *Lyngbya majuscula*, and is encoded by a mixed-polyketide nonribosomal-peptide assembly line (Chang et al., 2004).

The curacin PKS has an unusual terminal module. This module comprises an uncommon composition at its C terminal end which contains an ACP, a sulfotransferase (ST), and a TE. The ST domain catalyzes the transfer of a sulfonate group from adenosine 3′-phosphate 5′-phosphosulfate (PAPS) to a hydroxyl of an acceptor molecule (Figure 4-29) (Gehret et al., 2011; Gu et al., 2009).

An *In vitro* assay with heterologously expressed proteins revealed the mechanism of terminal olefin formation in biosynthesis of curacin A (Figure 4-29). It was shown, that insertion of a functional ST domain into the CurM termination module results in a unique process for polyketide chain release. A β-hydroxyl of the penultimate chain elongation intermediate converts into a β-sulfate which makes it a much more preferable leaving group. This chemical process afterward facilitates the following TE-mediated decarboxylative elimination. The crystal structure of this TE domain revealed how the familiar α/β-hydrolase architecture is adapted by this domain (Gu et al., 2009; Gehret et al., 2011). This enzyme superfamily is known for its catalytic promiscuity, and many members can perform more than one chemical reaction (C. Li et al., 2008). The same mechanism for terminal olefin formation is known for isoprenoid biosynthesis in the mevalonate pathway (Bonanno et al., 2001).

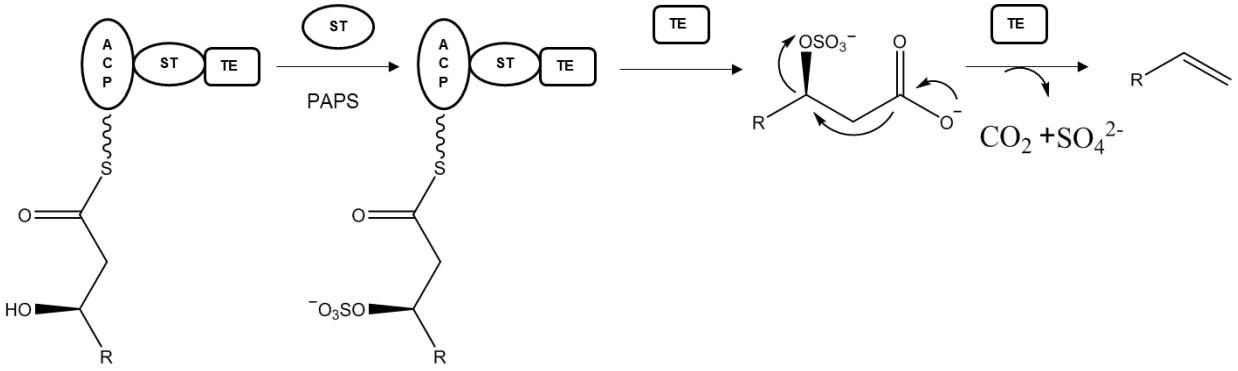


Figure 4-29. Decarboxylative elimination mechanisms and terminal olefin formation in the curacin A pathway (Gu et al., 2009).

ACP: acyl carrier protein, ST: sulfotransferase, TE: thioesterase, and PAPS: adenosine 3′-phosphate 5′-phosphosulfate.

Since no ST was encoded in the siphonazole biosynthetic gene cluster, the same proposed decarboxylation mechanisms as described for curacin A is not applicable for siphonazole.

In another study by Auldridge and coworkers on biosynthesis of methylketone, it was revealed that a hydrolase domain, methylketone synthase 1 (MKS1) can also catalyze the decarboxylation reaction. Methylketone-containing metabolites serve as toxic compounds against chewing insects and accumulate particularly in wild tomatoes. Biosynthesis of the methylketone in *Solanum habrochaites* f. *glabratum* starts with intermediates of *de novo* fatty acid synthesis. These fatty-acyl intermediates are later converted to the n-1 methylketone by the action of MKS2 and MKS1. The crystal structure of MKS1 revealed that this enzyme is an atypical member of the α/β-hydrolase superfamily. Notably, MKS2 a second locus in *S. habrochaites*, belongs to a family of hydrolase enzymes possessing the hot dog fold, which includes thioesterases such as 4-hydroxybenzoyl-CoA thioesterase and various putative fatty acyl hydrolases (Dillon & Bateman, 2004; Benning et al., 1998; Thoden et al., 2002). According to the described mechanism, MKS2 catalyzes in the first step the methylketone synthesis by cleavage of

the thioester linkage between the 3-keto acyl group and the ACP. Afterwards, MKS1 catalyzes the subsequent decarboxylation which results in 3-keto acid production (Figure 4-30) (G. Yu et al., 2010).

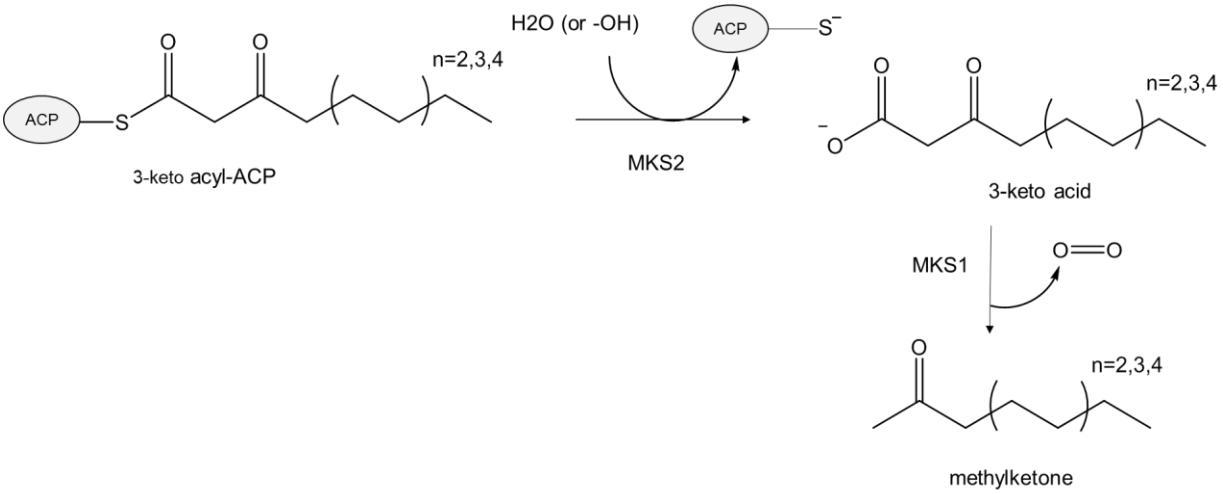


Figure 4-30. General reaction scheme for the production of methylketones.
MKS2 catalyzes the first step of methylketone synthesis, cleavage of the thioester linkage between the 3-keto acyl group and the ACP. Afterward, MKS1 catalyzes the subsequent decarboxylation of the resulting 3-keto acid.

The catalytic triad of the α/β-hydrolase consists of a nucleophile, a histidine and an acidic residue. The nucleophile residue is represented by serine which is highly conserved within this family (Kourist et al., 2010). Sequence comparisons revealed the MKS1 catalytic triad, Ala-His-Asn, as divergent to the traditional α/β-hydrolase triad, Ser-His-Asp. Thus, the highly conserved serine is replaced against an alanine in MKS1 catalytic triad.

In SphH the conserved serine residue is substituted against a cysteine moiety, but the histidine and an acidic residue are still present. The previous described model can be proposed for the terminal alkene formation in siphonazole. SphH could work as a decarboxylating enzyme, but still in this case the lack of a dehydration step has to be noted.

Terminal alkene formation in biosynthesis of tautomycetin, a selective protein phosphatase (PP) inhibitor from *Streptomyces* sp. CK4412 shows that presence of a dehydration step before decarboxylation is critical. In this gene cluster *ttnD* and *ttnF* encode putative decarboxylase and dehydratase enzymes, respectively. A study in which these enzymes were inactivated showed that TtnD apparently cannot catalyze decarboxylation in the absence of dehydratase TtnF (W. Li et al., 2008; Luo et al., 2010).

A DH domain is encoded within the *sphH* gene of the siphonazole cluster (Figure 4-10), but the essential and conserved histidine moiety is replaced with tyrosine. Hence, it can be assumed that this enzyme is inactive. The dehydration and the decarboxylation of the last incorporated acetate of siphonazole biosynthesis may thus be due to the last module, i.e. SphH, or SphJ alone, or even by cooperation of these two enzymes.

To clarify this, the heterologously expressed proteins of the putative hydrolase together with the adjacent encoded ACP (SphH Hyd-ACP), as well as of the TE SphJ were used in the assays to test their catalytical functions.

Expression of the SphH didomain Hyd-ACP in *E.coli* BAP1 cells resulted in soluble protein (Figure 4-17) which was utilized for PPant ejection assay. Applying the PPant ejection assay made it possible to analyse the molecule bound to the PPant arm. Therefore, depending on the nature of the loaded substrate on the ACP, after incubation of the model substrate with the purified didomain Hyd-ACP, the catalytical function of the putative hydrolase could be deduced (Figure 4-20 & 4-21). However, several efforts for detection of the expected masses failed. This can be caused by the big size of the applied didomian since the other experiments in the group working with the single ACP domain which had a lower molecular weight, were successful (Lohr et al. 2013). Therefore, heterologous expression of the single ACP domain of the SphH module and

performing the same can be helpful. To identify the catalytic role of the putative hydrolase, the constructed vector containing the encoding sequence of the hydrolase and ACP was successfully expressed in *E.coli* BL21 cells. The expression of the didomain in this strain results in inactivation of the ACP since this strain is not able to perform the co-expression phosphopantetheine transferase as well as the desired protein. By using both proteins in a reaction mixture, in the first step the substrate can be loaded on the phosphopantetheine arm of the single ACP and later the function of the putative hydrolase can be investigated.

Investigation of the off-loaded molecule was also performed in this study. It was assumed that the off loaded product of trans-1,3-pentadiene could be detactable either in chemical hydrolysis of the product from the didomain Hyd-ACP, or by using the catalytic activity of the TE SphJ.

The extracted products from the reaction mixtures containing either both or just the TE as well as the trans-1,3-pentadiene (from the reactions containing the proteins or without) was directly injected into the GC-MS instrument.

From the reaction mixture containing the authentic standard without protein a signal was observed using GC-MS (Figure 4-23). However, this was never observed in the main reaction mixture, and not even in the reaction containing the standard and protein. The disappearance of the free trans-1,3-pentadiene can be explained by binding of the released product to the protein, which makes it impossible to detect it with this method.

Anyway, since the final off-loaded product seems to be volatile it is worth to repeat the last experiment using an automated headspace solid phase microextraction method followed by GC-MS. This method was successfully used to evaluate and compare the *in vitro* production of microbial volatile organic compounds (Van Lancker et al., 2008). This sensitive method was well established in the laboratories of the Dickschat group at the Kekulé-Institut for Organic Chemistry and Biochemistry (University of Bonn, Germany). Therefore, the detection of the trans-1,3-pentadiene is currently performed in collaboration. This should on the long run enable to deduce the mechanism of decarboxylation and dehydration in the siphonazole biosynthetic process.

4.2.4.3 Origin of the starter unit in siphonazole biosynthesis

Feeding experiment with ^{13}C-labeled precursors revealed that the aromatic residue (protocatechuic acid) of the starter unit is derived from chorismic acid via the shikimate pathway (Nett. PhD thesis). On the other hand an experiments with the *O*-methyltransferase SphB, encoded upstream of the initiation module, showed that the enzyme readily methylates protocatechuic acid (Schäberle et al., 2013). Chorismate is the end-product of the seven metabolic steps of the shikimate pathway. The shikimate pathway starts with condensation of PEP and E4P, and results in production of aromatic amino acids tyrosine, tryptophan and phenylalanine as well as the aromatic building blocks for a number of secondary metabolites including enterobactin, pyochelin and yersiniabactin (Herrmann & Weaver, 1999; Van Lanen et al., 2008).

Chorismate lyase can transform chorismate into 4-hydroxybenzoate and pyruvate. 4-hydroxybenzoate can be converted to protocatechuic acid through hydroxybenzoate hydroxylase (PHBH; E.C. 1.14.13.2) (Siebert et al., 1996; Bertani et al., 2001). The conversion of chorismic acid to protocatechuic acid was observed in biosynthesis of phenazines. The phenazines number about 80 biologically active (antibacterial, antifungal, antiviral, and antitumor) secondary metabolites which are mainly synthesized by *Pseudomonas* and *Streptomyces* species (McDonald et al., 2001).

McDonald and coworkers (McDonald et al., 2001) have shown that incubation of chorismic acid with cell-free extracts from *E. coli* clones, which are overexpressing *phz* genes, results in conversion of the introduced substrate to the product protocatechuic acid. Furthermore, protocatechuic acid is a major metabolite in the degradation of aromatic compounds and can be directly generated from 4-hydroxybenzoic acid by *p*-hydroxybenzoate hydroxylase.

The DAHP synthase is the first incorporated enzyme of the shikimate pathway. SphI which is encoded near the downstream end of the siphonazole biosynthetic gene cluster shows the highest homology to DAHP synthases (Figure 4-10). Homologues of *dahp* synthase have already been discovered in the biosynthetic gene clusters of secondary metabolites, i.e. rifamycin, chloramphenicol and balhimycin (Floss & Yu, 1999; Shawky

et al., 2007; He et al., 2001). Therefore, the same function can be assumed for SphI in the biosynthesis of siphonazole. Furthermore, analysis of the genome sequence revealed that orf1 located upstream of the siphonazole cluster, is encoding a PEP synthase. Therefore, this isoenzyme can supply the DAHP synthase with PEP, one of its required substrates. Protocatechuic acid, synthesized via the shikimate pathway, was hypothesized to be the starter unit of siphonazole biosynthesis. This was proven by performing the ATP-PP_i exchange assay which showed the specificity of the A domain in the invitation module towards protocatechuic acid (Höver. unpublished result). The here performed *in vitro* assay with the DAHP synthase proved the involvement of SphI in precursor formation for siphonazole biosynthesis. The optimal conditions as well as the kinetic parameters for this enzyme were determined. The enzyme reached its maximum activity at 30°C and in the neutral pH (pH 7) (See 4.2.3.3 and figure 4-27). This results could be expected since it agrees with natural growing conditions of the original host *Herpetosiphon* sp. 060 which was isolated from a mud sample of the intertidal region (Nett & König. 2007). The metal dependence of SphI was experimentaly proven, since EDTA-treatment of the enzyme resulted in no activity, while the enzyme was always active in presence of metal ions. The metal requirements of the *Herpetosiphon* sp. 060 DAHP synthase are similar to those reported for DAHP synthase from other microbial sources. Activity of DAHP synthases from *T. maritima* (Wu et al., 2003)*, A. nidulans* (Hartmann et al., 2001) and the two DAHP synthases, NCgl0950 (AroF) and NCgl2098 (AroG) of *C. glutamicum* were significantly abolished when these enzymes were treated with EDTA (Y. J. Liu et al., 2008).

The highest SphI activity was observed in presence of Mn^{2+} and Cu^{2+}, whereby the activity was quite low in the presence of Zn^{2+} (Table 4-4). Summarizing the *in vitro* assays, as well as the presence of the conserved metal binding sequence motif I in SphI, it can be concluded that this enzyme is indeed a metalo-enzyme (Figure 4-25).

Comparison of the maximum velocities of SphI ($0.19 \times 10^{-3} \pm 0.1 \times 10^{-3}$ U/mg^{-1}) and Amb7 (V_{max} 0.3725 ± 0.083 U/mg^{-1}) indicates that Amb7 is about twice as active as SphI (See 4.2.3.6 and figure 4-8). The K_m values obtained were 0.6993± 0.3746 mM and

0.3779 ± 0.0208 mM for Amb7 and SphI, respectively, indicating that the SphI exhibits much higher substrate affinity than Amb7.

4.3 Introduction into the corallopyronin A biosynthetic gene cluster and biosynthetic hypothesis

Corallopyronin A (Figure 1-13), a myxobacterial antibiotic is synthesized by a hybrid of PKS and NRPS. The corresponding gene cluster was already identified (Erol et al, 2010) (Figure 4-31).

The molecule is synthesized from two separate chains (eastern and western chain), which are finally intermolecularly connected to form the precursor, i.e. precorallopyronin A, of the active compound. The starter molecule of the eastern chain is hypothesized to be carbonic acid, which becomes methylated and subsequently is elongated by a glycine moiety and three acetate units. In that way, the most unusual vinyl carbamate functionality is formed by the mixed NRPS/PKS system involved. This assumption was based on the domain order of the biosynthetic enzymes and on the feeding experiments with labeled precursors, i.e. acetate, bicarbonate, glycine, and S-adenosylmethionine (SAM). The incorporation of these precursors into the final molecule was clearly proven. Formation of the methyl ester was suggested to involve the *O*-methyltransferase CorH (Erol et al., 2010). We investigated in the present study this putative *O*-methyltransferase and the attachment of carbonic acid methyl ester at the ACP, i.e. ACP1 of the loading module of the corallopyronin A assembly line. The western chain is solely PKS derived and, with the exception of the two methyl groups in the β position, which are introduced by a specific β-branching cassette encoded in the gene cluster (Erol et al., 2010), its assembly follows the classical colinearity rule (Piel, 2010). However, the enzymes involved in the assembly of the western chain have been hypothesized to show a certain degree of flexibility, since the derivative corallopyronin B was detected in the fermentation broth (Schäberle et al., 2013), whereby the latter might be based on a different building block incorporated, i.e. propionyl-CoA (propionyl coenzyme A) instead of acetyl-CoA as the starter unit.

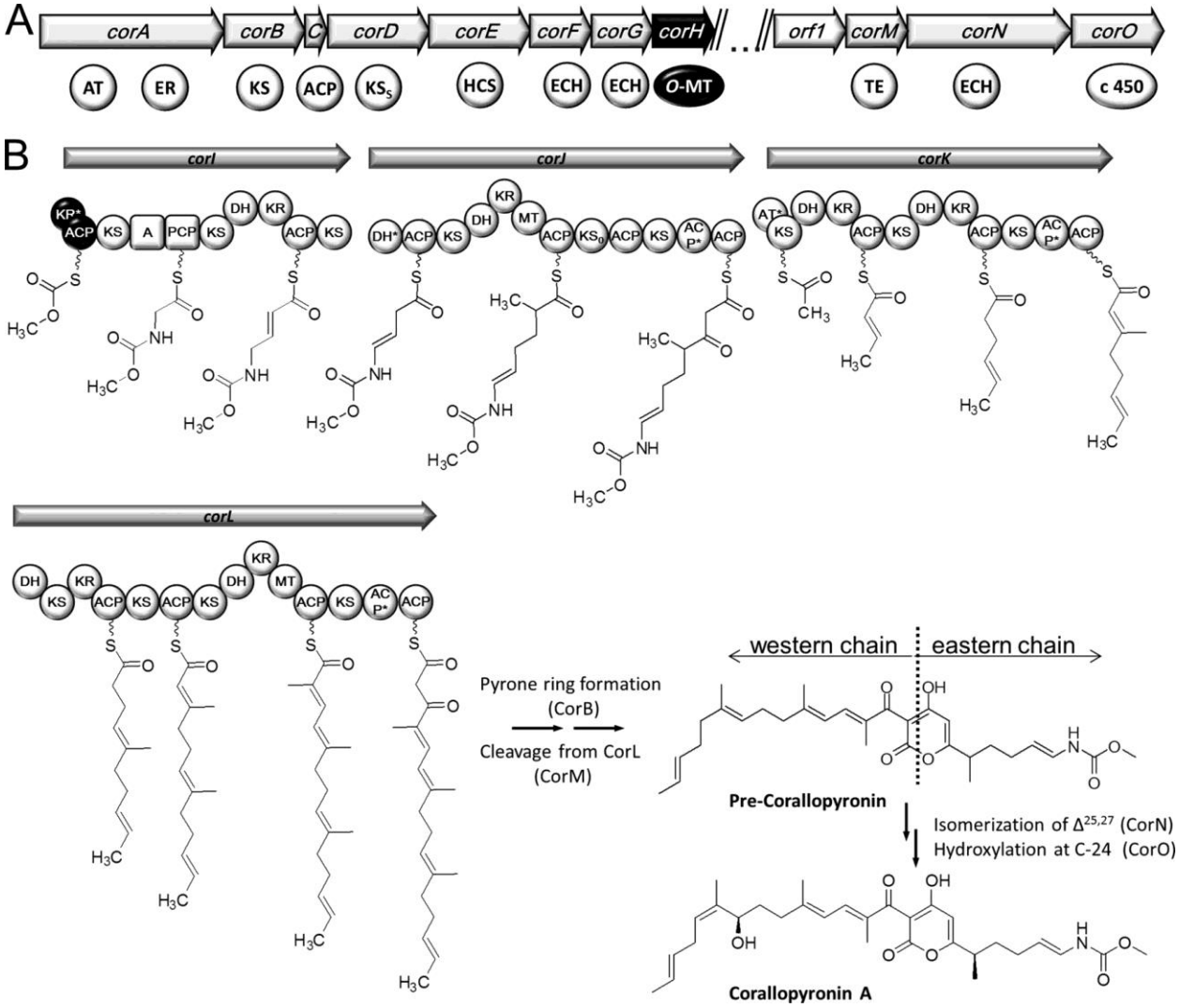


Figure 4-31. Corallopyronin A biosynthetic gene cluster and biosynthetic hypothesis. (A) Corallopyronin A biosynthetic gene cluster drawn to scale. The smaller genes are depicted, whereby the core PKS genes are shown below. CorA, *trans*-acyltransferase/enoylreductase; CorB, (unusual) ketosynthase; CorC, acyl carrier protein; CorD, (unusual) ketosynthase; CorE, hydroxymethylglutaryl (HMG)-CoA synthase; CorF and CorG, enoyl CoA-hydratase/isomerase; CorH, *O*-methyltransferase, which was used in the present study; CorI, mixed nonribosomal peptide synthetase/polyketide synthase (NRPS/PKS); CorJ-L, PKS; CorM, thioesterase; CorN,

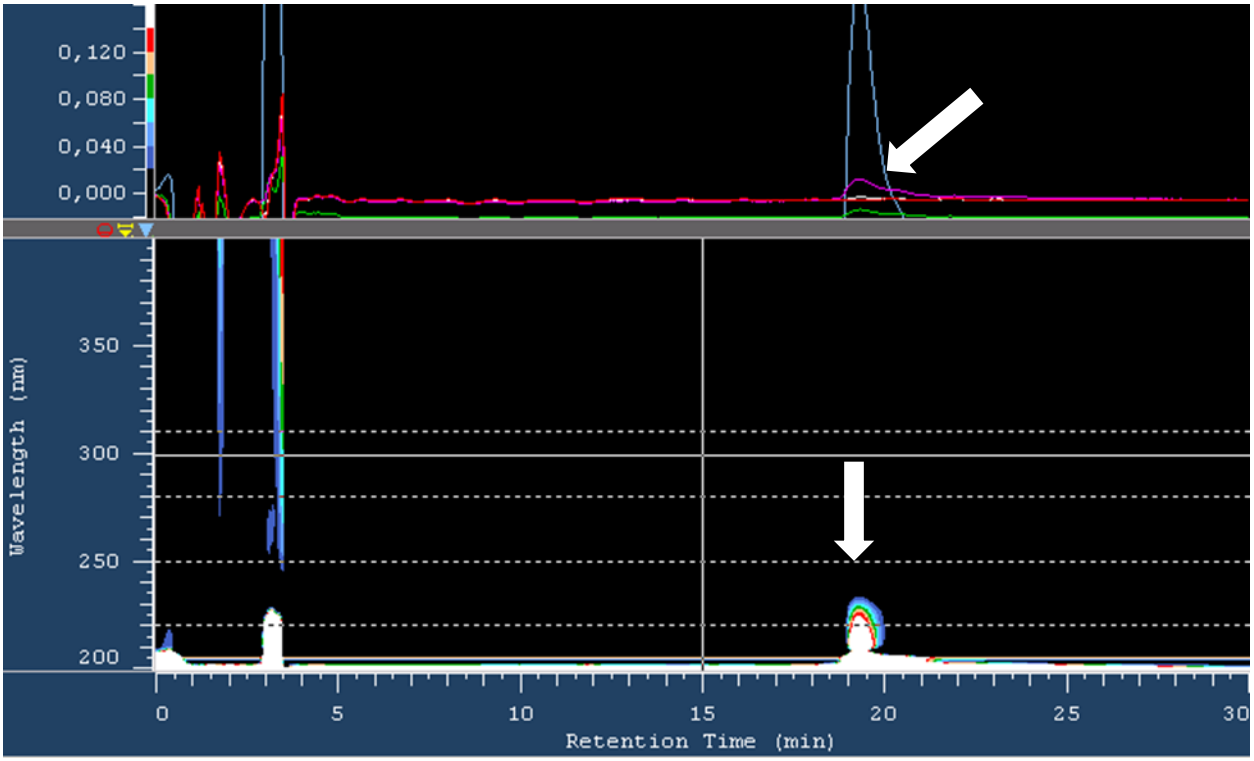
isomerase (ECH); CorO, cytochrome P450 protein. (B) Biosynthetic hypothesis. The arrows represent the genes; below, the corresponding protein domains are depicted. A, adenylation domain; ACP, acyl carrier protein; DH, dehydratase domain; KR, keto reductase domain; KS, keto synthase domain; KS0, inactive KS; MT, methyl transferase domain. The AT*, ACP*, and KR* domains are degraded. DH* marks a shift domain. The enzyme(domain)s heterologously expressed in the present study are shown in black with white characters (Schäberle et al., 2013).

This study focused on the formation of vinyl carbamate residue, situated at the terminal part of the eastern chain in the final molecule. Therefore, the catalytical function of the loading module in CorI and of the *O*-methyltransferase CorH, were investigated in *in vitro* assays described in 4.3.1 and 4.3.2.

4.3.1 Loading of methylated hydrogen carbonate onto ACP1

Biosynthesis of the eastern chain of corallopyronin A (Figure 4-31) was predicted to use carbonic acid or its monomethyl ester as the starter molecule (Erol et al, 2010). To experimentally verify this hypothesis, the expression of the first acyl carrier protein (ACP1) domain of the biosynthetic gene cluster was projected. Therefore, the 5´ part of the *corI* gene was amplified by PCR with specific primers, yielding a PCR product of 1432 bp. The latter was subsequently cloned into the pET151 expression vector, which resulted in the pET151-ACP1 construct (Method 3.12.15). This vector was transferred to *E. coli* BaAP1 cells. After expression and purification, the ACP1 *holo*-protein was incubated with freshly prepared crude protein extract from strain *C. coralloides* B035, which was grown for 48 h, a time frame in which corallopyronin A production was always observed. Incubation was performed for 70 min at 30°C. Longer incubation times resulted in protein degradation. The ACP1 protein treated in that way was desalted and purified by HPLC (Method 3.17) (Figure 4-32).

(A)



(B)

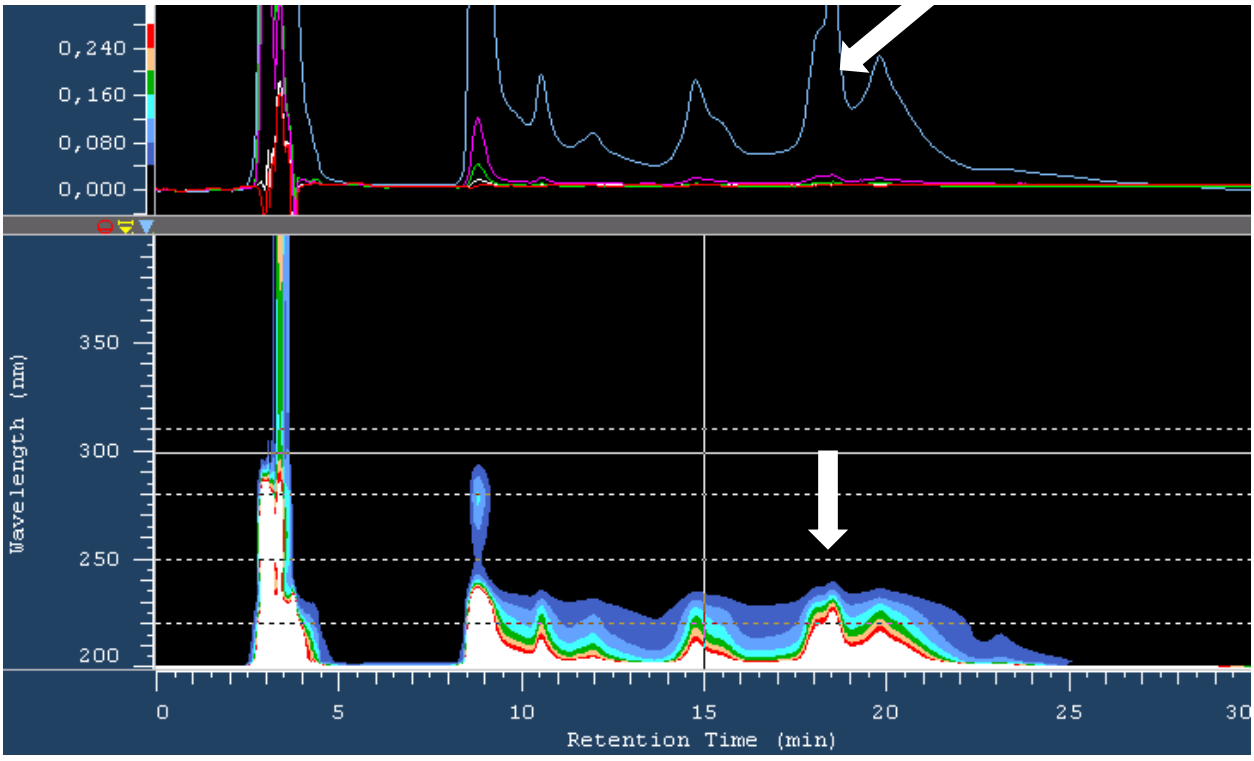


Figure 4-32. HPLC chromatograms of the purification and desalting of the ACP1 protein of the corallopyronin A biosynthetic assembly line.

(A) Purified ACP1; (B) after treatment with crude cell extract from *C. coralloides* B035, the white arrow points towards the ACP1 containing fraction.

After the ACP1-containing fraction was dried, analysis was performed by electrospray ionization-mass spectrometry. Applying the PPant ejection assay (Meluzzi et al., 2008) made it possible to identify substrates loaded onto *holo*-ACP1. Collision energy in the linear ion trap applied to such a *holo*-protein resulted in the ejection of the PPant residue. If a substrate is tethered to this PPant arm, ejection ions will consist of PPant+substrate (Meluzzi et al., 2008; Dorrestein et al., 2006). A further round of fragmentation clarifies whether the isolated ejection ion is indeed a PPant-containing molecule, since specific signature ions, corresponding to the PPant residue, can be used for the identification. For the ACP1 *holo*-protein incubated with crude extract from *C. coralloides* B035, an ejection ion (MS^2) was identified, whose *m/z* (319.13, isolation width 1) corresponded to methylated hydrogen carbonate tethered to the PPant arm (calculated *m/z*, 319.13) (Figure 4-33).

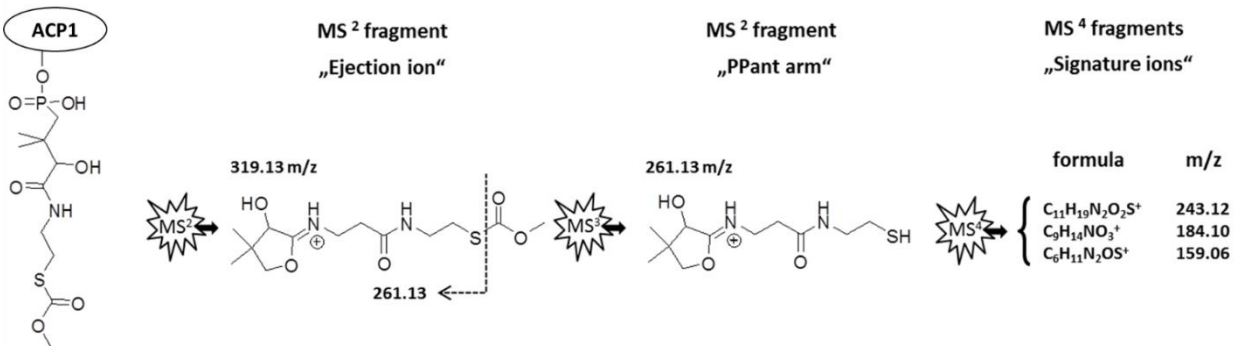


Figure 4-33. Phosphopantetheine (PPant) ejection assay with ACP1 of the corallopyronin

A biosynthetic enzymes, incubated with *C. coralloides* B035 crude extract. The PPant arm plus attached substrate, *i.e.* carbonic methyl ester is liberated from the *holo*-protein

ACP1 by applying collision energy, and resulted in an ejection ion. Further rounds of fragmentation (MS^3 and MS^4) revealed the diagnostic signature ions related to the PPant arm.

Fragmentation of the ion *m/z* 319.13 (MS^3) resulted in a free PPant arm (calculated *m/z*, 261.13; measured *m/z*, 261.13). To unambiguously verify the nature of the ejection ions, a further round of fragmentation (MS^4) was applied. The expected signature ions for the PPant arm were calculated to be *m/z* 243.12, 184.10, and 159.06 and thus matched the measured values of *m/z* 243.12, 184.10, and 159.06 (Figure 4-33 & 4-34).
This clearly proved the nature of the parent ejection ion as loaded onto the PPant arm, and evidence was gained that the predicted unusual building block is indeed methylated hydrogen carbonate, acting as the starter molecule by priming the loading module for corallopyronin A biosynthesis.

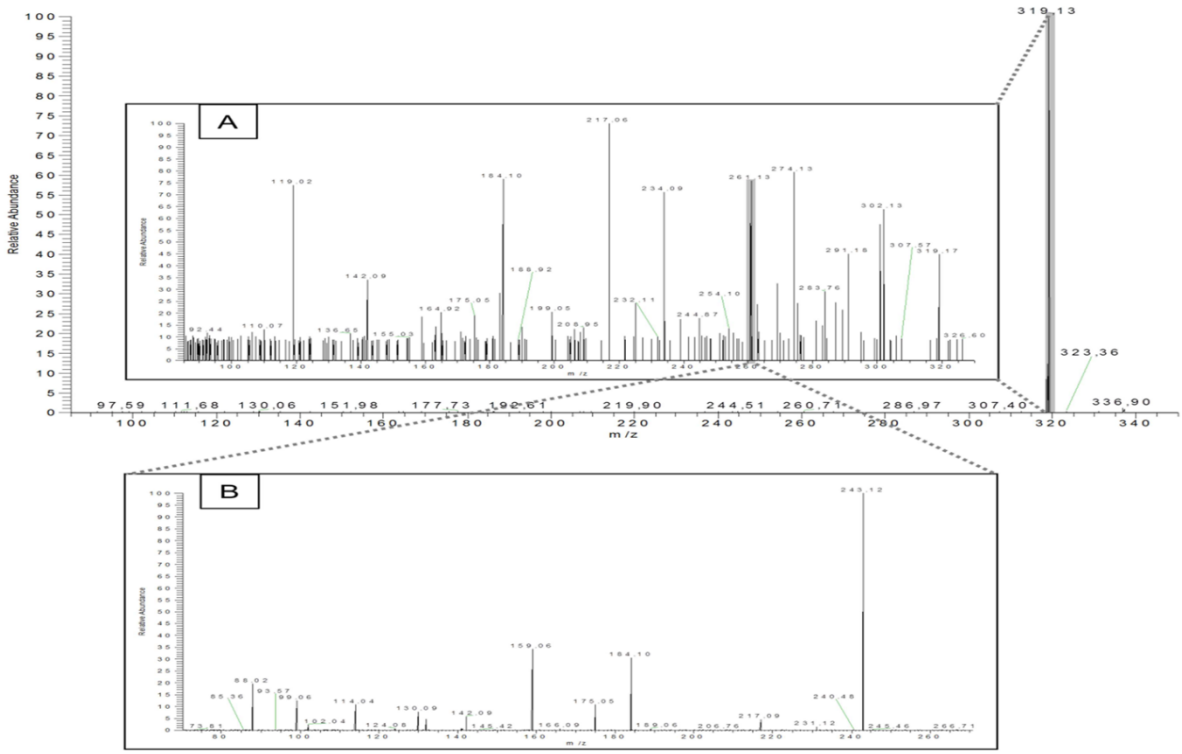


Figure 4-34. Fragmentation procedure of ions with *m/z* 319.13.

The ejection ion *m/z* 319.13 was isolated and fragmented, yielding inlet A. To prove the identity of the phosphopantetheine (PPant) arm the ions with *m/z* 261.13 were isolated and subjected to a further round of fragmentation, yielding inlet B. Inlet B shows the specific signature ions for a PPant arm (compare figure 4-33).

4.3.2 Analysis of the SAM-dependent O-methyltransferase CorH

The analysis of the ACP1 bound starter molecule of corallopyronin A biosynthesis indicated that the *O*-methylation reaction of the C1-building block takes place before the starter unit is tethered to the PKS/NRPS machinery, since un-methylated carbonic acid was not detected. The putative candidate for this methylation reaction was CorH (Figure 4-30), showing according to bioinformatics analysis similarity to *O*-methyltransferases (MTs). MTs are often found as domains within the PKS machinery and act on the growing intermediate, or are clustered with the PKS genes and code for enzymes catalyzing so-called post-PKS reactions on the molecule backbone built by the PKS core enzymes (Rix et al., 2002). The latter seemed to be the case for CorH whose gene *corH* is encoded directly upstream of *corI*. CorH, however seems to act on a precursor molecule prior polyketide assembly.

Comparison of the amino acid sequence encoded by *corH* with proteins found in data bases in a protein BLAST search revealed some sequence identity with putative SAM-dependent MTs from mycobacterial species (identity ~35%). MelK, an *O*-MT from *Melittangium lichenicola* (Müller & Müller, 2006; Inga Müller et al., 2006) showing 27% identity towards CorH, was the first hit in the group of the δ-proteobacteria, and is catalyzing an *O*-methylation in the biosynthesis of myxothiazol, also a myxobacterial PKS/NRPS-derived antibiotic.

The signature sequences for SAM binding were difficult to identify in CorH and the related sequences like in MelK. However, deeper analysis of the CorH sequence revealed conserved motifs that are predicted to contribute to SAM binding. In comparison with the biochemically investigated human brain protein phosphatase 2A

leucine carboxyl methyltransferase (LCMT) (De Baere et al., 1999), and with MelK, both MTs use SAM as co-factor, and the binding motifs I-III could be detected (Figure 4-35). These regions coincide with the conserved motifs identified and predicted to contribute to SAM binding (Kagan & Clarke, 1994). Thus, it can be expected that CorH, as MelK and LCMT, binds the cofactor SAM in a similar way as other MTs, as judged from the presence of the typical SAM binding fold.

```
CorH     82 QFLILGSGLD...102 VRYFEVE...132 VPAEY...156 TFILWEGNVFYL
Q9X7D5  103 QAVIVAAGLD...124 ATVFEID...156 VAADL...179 SAWSVEGLLPYL
LCMT1    93 QIVNLGAGMD...116 SKYFEVD...168 IGADL...193 TLLIAECVLVYM
MelK     94 QVVLLGAGYD...116 VRLFELD...148 VPINF...172 TFFLWEGVSMYL
              * : :.:* *          **::       :  :        :    *    *:
```

Figure 4-35. SAM binding motifs of CorH and selected homologues.

The Kagan and Clarke motifs (order: motif I, post I, motif II, and motif III) are shown. The always conserved Gly in position 89 (CorH numbering) is highlighted by a black box. Asterisks (*) and dots (**:** and **.**) mark homologous and similar residues. Amino acid sequences used for the alignment were derived from the following organisms (accession numbers in brackets): *Corallococcus coralloides* B035 (HM071004), *Mycobacterium leprae* (Q9X7D5), *Homo sapiens* LCMT1 (AAF18267), *Melittangium lichenicola* (AJ557546).

To support experimentally the proposed function of CorH as an *O*-MT involved in the formation of the starter, its heterologous expression in *E. coli* was projected. However, all attempts to express and purify CorH as a His-tag, His-Asn-tag, or Streptavidin-tag fusion protein failed, since the expression level was too low. For that reason, finally a synthetic gene was used, whose codon-usage was optimized for the heterologous host *E. coli*. Therefore, the tool GeneOptimizer® was employed, which is embedded in the pipeline for gene synthesis by GeneArt®. The codon-usage optimized gene was

constructed with introduced *Eco*RI and *Hin*dIII restriction sites. Using these restriction sites the gene was excised from the original plasmid, gel-purified, and subsequently ligated into the likewise restricted expression vector pET28, yielding plasmid pET28CorH* (Method 3.12.16).

A second obstacle which had to be overcome to biochemically characterize CorH was, that standard HPLC-based assays, like HPLC coupled with diode array detector (DAD), could not be used, since hydrogen carbonate shows no UV absorbance. Therefore, a newly developed assay, which is coupling the activity of *S*-adenosyl homocysteine (SAH) hydrolase to the SAM-dependent methylation reaction, was used (Schäberle et al., 2013). Applying this approach, the methylation activity could be determined indirectly by measuring the amount of SAM consumed. To avoid the inclusion of unspecific SAM-consumption through remaining activity from contaminating *E. coli* proteins additional controls had to be made. Indeed, in contrast to the negative controls without enzyme in which the methylation activity was zero, yielded control reactions with enzyme but without substrate detectable SAH. This value was subtracted to set the basal background activity to zero. In that way the remaining activity could then be assigned to CorH and the consumed substrate. Using this assay, CorH activity towards the expected substrate hydrogen carbonate was proven. To test the activity of CorH towards other substrates presenting a hydroxyl group glycerol and isovanillinic acid were applied. Only with glycerol, which presents three hydroxyl groups accessible for methylation, activity was detected. However, under the prevailing conditions the methylating activity was very low. After 26 h incubation at 30°C 25.7 µM SAM were used to methylate HCO_3^- and 19.5 µM to methylate glycerol, respectively. Further, the enzyme proved itself very heat-sensitive, and activity was detected at 30°C but unexpectedly no activity remained in samples incubated at 37°C.

4.3.3 Conclusion on the catalytical function of the CorI loading module and of the *O*- methyltransferase CorH in vinyl carbamate formation in corallopyronin A biosynthesis

Thus far, all isolated derivatives of corallopyronin A (Figure 1-13) and the structurally related myxopyronins (Figure 4-36) have shown no alteration on the eastern chain and start with the unusual vinyl carbamate functionality. Sucipto et al. reported a gene knockout experiment for the myxopyronin producer, the genetically accessible strain *Myxococcus fulvus* Mxf50. In this strain, the *corH* homologue *myxH* was deleted, resulting in a complete abolishment of myxopyronin A production (Sucipto et al., 2013), thereby proving this methyltransferase, i.e. MyxH and CorH, to be enzymes essential for biosynthesis of these antibiotics.

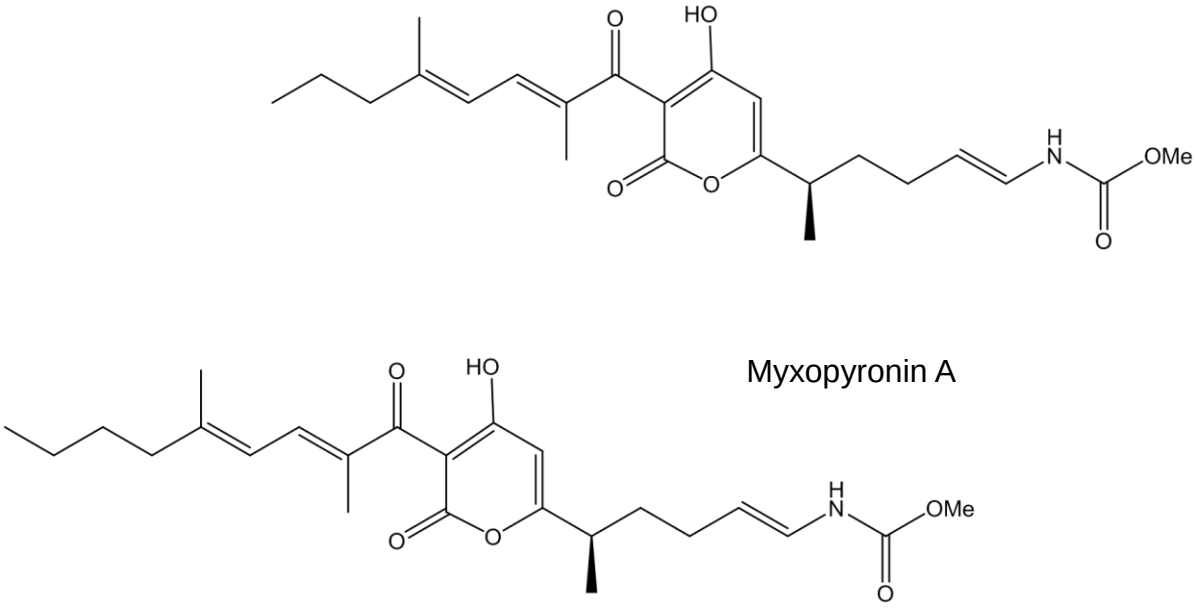


Figure 4-36. Molecular structures of myxopyronin A and myxopyronin B.

Feeding studies with labeled precursors for corallopyronin A biosynthesis revealed SAM and sodium bicarbonate as constituent parts of this moiety (Erol et al., 2010). Therefore, the biosynthetic hypothesis was that carbonic acid or, more plausibly, its more stable monomethyl ester is the first building block loaded onto the PKS/NRPS assembly line. Usually, building blocks have to be present as coenzyme A (CoA) esters and are loaded by the catalytic activity of ATs. However, no such enzyme is encoded within the biosynthetic gene cluster for corallopyronin A, and such an enzyme encoded elsewhere in the *C. coralloides* B035 genome might be responsible. Another possibility is the self-loading of the substrate, as has already been described for other ACPs (Hitchman et al., 1998). PKS substrates are in general activated as CoA or, in the case of (amino) acids, activation is achieved by phosphate groups that enable subsequent loading onto the NRPS carrier protein domains. Such a mechanism can be expected for the incorporation of hydrogen carbonate, whereby we suspect that first a phosphate group from ATP may be transferred to HCO_3^- to give carbonyl phosphate. This would mean that HCO_3^- connected with an activating group rather than the free molecule would represent the natural substrate, accessible to *O*-methylation by CorH. That could explain the low activity of CorH in the performed assays, whereby sodium bicarbonate which dissolves into free HCO_3^- was used. The instability of CorH at increased temperatures, i.e. activity was unexpectedly completely abolished at 37°C, reflects the growth optimum of the antibiotic-producing strain at 30°C. This is in contrast to other investigated MTs that show higher temperature optima before an activity decrease takes place, e.g. the maximal activities of MycE and MycF, two MTs involved in the biosynthesis of mycinamicin macrolid antibiotics, were observed at 50 and 37°C, respectively (S. Li et al., 2009). The successful methylation of HCO_3^- by CorH through SAM consumption is in agreement with feeding studies performed earlier, which clearly showed SAM to be the methyl group donor (Erol et al., 2010). Thus, here the assumed involvement of CorH in allocation of a precursor for corallopyronin A biosynthesis was verified. It can further be assumed that binding to an activating group such as, e.g. phosphate, and methylation are early events to get a more stable precursor available for the biosynthesis (Figure 4-36), since free HCO_3^- is in equilibrium with CO_2, which is volatile.

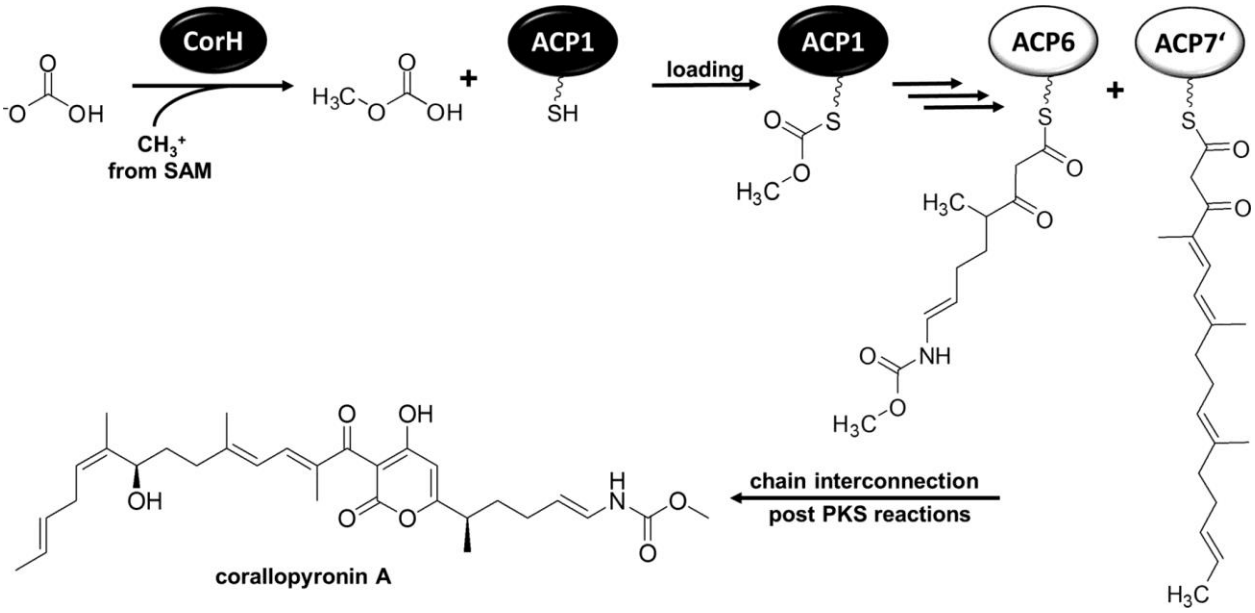
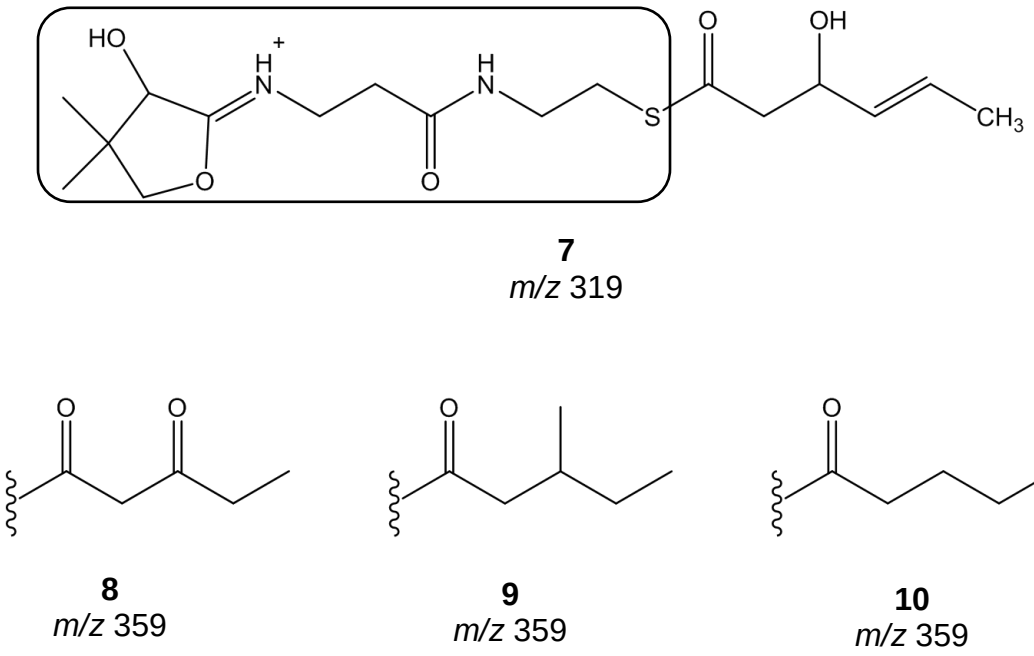


Figure 4-36. Highlighted steps in the biosynthesis of corallopyronin A and of its precursor by CorH. CorH catalyzes the methylation of hydrogen carbonate by SAM, yielding methylated hydrogen carbonate. This is the precursor which gets in the following loaded to the carrier protein of the loading module (ACP1) of the mixed *trans*-PKS/NRPS enzymatic machinery of corallopyronin A biosynthesis. After tethering of several more building blocks, e.g. acetate, glycine, and SAM, the eastern chain is completed at the carrier protein of module 6 (ACP6). By head-to-head condensation with the western chain, which is bound to the carrier protein of module 7´ (ACP7´), the corallopyronin A backbone is build. Further processing (post-PKS reactions) finally yields corallopyronin A.

Experiments with ACP1 showed that the concentration and stability of the precursor, which initiates corallopyronin A biosynthesis, is indeed high enough in *C. coralloides* B035 cells to enable loading of the carrier protein by incubation with cell lysate.

Whether the precursor is methylated hydrogen carbonate itself or its phosphate ester cannot be judged from our experiments. The first carrier protein ACP1 of corallopyronin A biosynthesis is either loaded by a transferase-like enzyme or by self-loading, whereby

in each case the precursor is specifically picked out of the cell lysate. *Trans*-Acyl transferases (*trans*-ATs) show a preference for specific activated building blocks which they load to the corresponding ACPs. By specific interaction of ACPs and ATs, the recognition and loading of the correct substrate is assured (Musiol et al., 2011). In the corallopyronin A gene cluster, only one *trans*-AT is encoded, i.e., in *corA*, that is suspected to be responsible for the loading of the malonyl-CoA units to the respective modules. Thus, it is more likely that CorA is not involved in the loading of the unusual starter unit; rather, a self-loading mechanism might be the case. For carrier domains of NRPSs, it has been shown that they are able to select the correct substrate from complex mixtures (Dorrestein et al., 2006). However, to a small extent, additional unspecific loading of ACPs in NRPSs has been demonstrated to occur within such experiments (Musiol et al., 2011). In addition to these observations, unspecific loading of ubiquitous CoA derivatives was found for expressed *holo*-carrier proteins (Lohr et al., 2013). In these cases, ejection ions with an *m/z* of 303 were observed (Lohr et al., 2013), corresponding to a loaded acetyl moiety. In the case of ACP1, no such acetylated variant was detectable. However, by scanning for a PPant signature, a second ejection ion with *m/z* 359 (Isolation width 1) could be isolated, representing most probably an unspecific loading with a CoA derivative arising from fatty acid metabolism (Figure 4-37).



7
m/z 319

8
m/z 359

9
m/z 359

10
m/z 359

Figure 4-37. Some possible ejection ions of molecules bound to ACP1 of corallopyronin A biosynthesis. The box marks the phosphopantetheine (PPant) part to which the building blocks are tethered. The calculated *m/z* values of singly charged ions [M+H$^+$] are given. **7** ejection ion of methylated hydrogen carbonate tethered to PPant, **8**-**10** molecules yielding an *m/z* 359, which might be tethered to PPant.

Usually, such wrongly loaded building blocks do not appear in the final products of PKS/NRPS assembly lines, since the downstream following domains fulfill gatekeeper functions and terminate the process. The assembling enzymes for the western chain, however seem to accept to a small extent different starter molecules, resulting in two derivatives with an altered chain length, namely, corallopyronins A and B. Both show good antimicrobial activities, whereby the activity of the major metabolite corallopyronin A outcompetes corallopyronin B (MICs of 0.097 µg/ml versus 0.39 µg/ml toward *Staphylococcus aureus*) (Irschik et al., 1985). It can be assumed that the basis for this difference in the MICs is that corallopyronin A fits better into the binding pocket of the antibiotic within the switch region of the RNAP. This would be in accordance with the

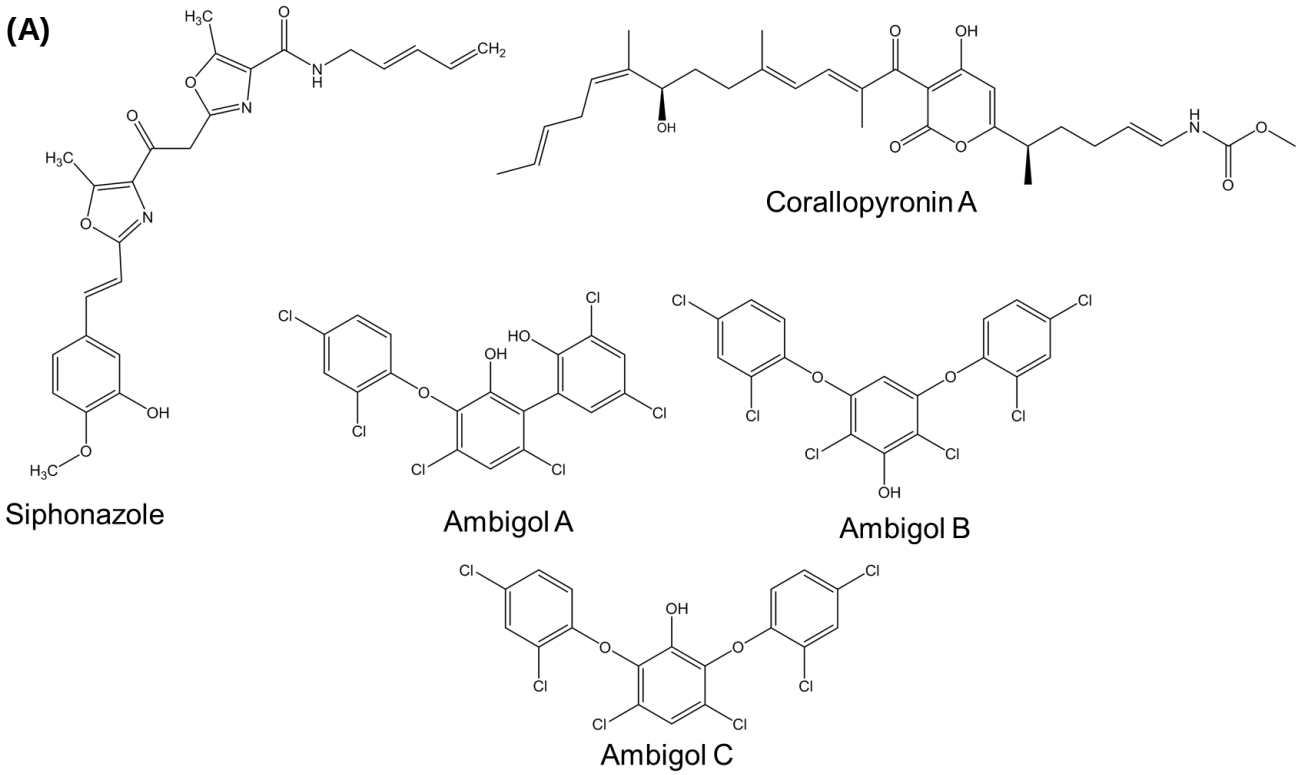
fact that the antibacterial activity increases from the structurally comparable myxopyronin A to myxopyronin B to corallopyronin A with the increasing length of the western chain (MICs of 1, 0.3, and 0.097 µg/ml, respectively, toward *S. aureus*) (Irschik et al., 1985). It seems as though the available space within the hydrophobic pocket is optimally occupied by corallopyronin A (Schiefer et al., 2012), and even slight elongation by one carbon atom, as in corallopyronin B, results in a repulsion reduced fit to the target structure.

5 Summary

Natural products are most important in drug discovery and therapy, e.g. more than 80% of antibiotics are derived from microorganisms. Antimicrobial resistance is a global threat and thus an urgent need for new antimicrobial compounds is evident. Gliding bacteria are a unique group of natural product producers, hereto hardly investigated for novel drug-like molecules.

The present study focussed on gliding bacteria-derived secondary metabolites. These include the highly halogenated and aromatic ambigols from the cyanobacterium *Fischerella ambigua*, siphonazole from *Herpetosiphon* sp. 060, belonging to the Chloroflexi, and corallopyronin A produced by the myxobacterium *Corallococcus coralloides* B035. The main aim of this investigation was to shed light on the biosynthetic process, especially to determine the starter units for the formation of these metabolites. In addition, in the case of siphonazole biosynthesis, a complete heterologous expression of the siphonazole biosynthetic gene cluster was attempted and the enzymes involved in the synthesis of the terminal alkene moiety were probed.

Concerning the cyanobacterial ambigols, 4-hydroxy benzoic acid (4-HBA) was postulated to serve as the precursor in the biosynthetic process. The latter was suggested to be generated from chorismate via the shikimic acid pathway. In the putative ambigol biosynthetic gene cluster the enzyme Amb7 is encoded, which shows high identity to 3-deoxy-D-arabinoheptulosonate7-phosphate (DAHP) synthases. Therefore, this enzyme was assumed to provide the chorismate precursor for 4-HBA production by condensation of phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P). To test this hypothesis, the enzyme Amb7 was heterologously expressed and *in vitro* studies with purified Amb7 proved this enzyme as a DAHP synthase. In further experiments the parameters for optimal activity of this enzyme were determined, including temperature and pH optimum, as well as metal requirements.



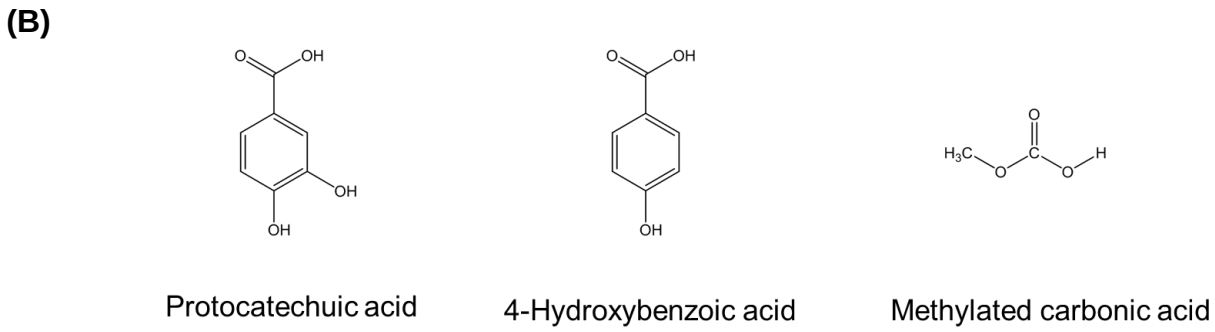


Figure 1. (A) Structures of ambigols, siphonazole and corallopyronin A. (B) suggested starter units in their biosynthesis.

In the biosynthesis of siphonazole, also a shikimic acid pathway-derived precursor is involved, as proven by feeding studies with ^{13}C-labeled hydroxybenzoic acid in the PhD thesis of Markus Nett (2007). In the putative siphonazole biosynthetic gene cluster the DAHP synthase SphI is encoded, and suggested to be involved in chorismate formation. SphI was expressed, purified and its catalytical activity examined in the same way as described for Amb7, using spectrophotometric assays. Both enzymes were shown to be metalo enzymes, whereby their optimal temperature (30°C) and pH (i.e. pH 7) reflects the natural growth conditions of the respective bacterial species.

A further series of experiments targeted the detailed characterization of the terminal alkene formation in siphonazole biosynthesis. To obtain the unusual alkene terminus, a decarboxylation and a dehydration reaction was suggested to occur. In the 3' region of the corresponding gene cluster a putative hydrolase (Hyd) domain is encoded in the *sphH* gene, and additionally the thioesterase (TE) SphJ. These represent good candidates for the proposed reactions. Therefore, the putative hydrolase domain of SphH together with the adjacent encoded acyl carrier protein (ACP) (SphH Hyd-ACP), as well as the TE SphJ were expressed in soluble form in *E. coli* BAP1 cells. Subsequently the catalytical functions of these enzymes were tested. To analyze the didomain Hyd-ACP of SphH, a substrate mimic was chemically synthesized using commercially available trans-3-hydroxyhex-4-enoic acid, which was activated by N-acetylcysteamine (SNAC)-coupling. Unfortunately the ACP-tethered 3-hydroxyhex-4-enoic acid was not converted by Hyd and TE, as evidenced by MS measurements. Therefore, using this approach currently no further insights into the catalytic role of SphH Hyd-ACP as well as TE could be obtained.

In a second approach, analyses of the ejected molecule, i.e. trans-1,3-pentadiene by GC-MS was attempted. Therefore, the off-loaded product obtained either by chemically hydrolyzing the thio-ester using KOH or by using the heterologously expressed TE in the reaction mixture, was extracted or analyzed. Although the ejection of the extracted standard from the reaction mixture without the protein solution resulted in the expected GC peaks with a molecular weight of 68 Da, proof of either (E)-penta-1,3-diene or

(Z)-penta-1,3-diene production in the reaction mixtures containing the heterologous protein solution failed.

Much effort went into the heterologous expression of the siphonazole gene cluster in Bacilli, as well as Streptomycetes. To this purpose *Bacillus amyloliquefaciens* FZb42 was chosen as the expression host. To integrate the construct harboring the complete biosynthetic gene cluster into the *Bacillus* genome, the gene *baeJ* served as target sequence for homologous recombination. To establish a suitable process to transform *B. amyloliquefaciens* first a smaller vector (pGF27+*baeJ*) containing the *baej* gene was successfully transfered. The same conditions were applied for transferring the construct (pCC1FOS-siphonazole) containing the complete siphonazole gene cluster as well as a copy of *baeJ* gene. However, no recombinant colony was obtained. This might be due to the size of the construct, i.e. ≥70 kb which can cause stability problems. Thus, the reconstitution of the gene cluster in a bacterial artificial chromosome (BAC) vector was programmed. However, all attempts of cloning the siphonazole gene cluster in the BAC vector failed.

Corallopyronin A is a promising *in vivo* active antibiotic, currently undergoing preclinical evaluation. This myxobacterial compound interferes with a newly identified drug target site, i.e. the switch region of the bacterial DNA-dependent RNA-polymerase (RNAP). Since this target site differs from that of known RNAP inhibitors such as the rifamycins, corallopyronin A shows no cross-resistance with other antibacterial agents. Corallopyronin A is a polyketide synthase- and nonribosomal peptide synthetase-derived molecule whose structure and biosynthesis is distinguished by several peculiarities, such as the unusual vinyl carbamate functionality whose formation involves carbonic acid as an unprecedented C1-starter unit. Using *in vitro* experiments the nature of this starter molecule was revealed to be the methyl ester of carbonic acid. Biochemical investigations showed that methylation of carbonic acid is performed by the *O*-methyltransferase CorH. These experiments shed light on the biosynthesis of the Eastern chain of α-pyrone antibiotics such as corallopyronin A. In conclusion, the DAHP synthases providing the starter unit in the biosynthesis of ambigols and siphonazole were investigated. In addition, *in vitro* experiments with CorH and CorI shed light on the

biosynthesis of the eastern chain of corallopyronin A. Further experiments toward a detailed understanding of siphonazole biosynthesis, either via heterologous expression of the complete gene cluster or through expression of individual protein(domain)s of the terminal part were unfruitful up to now. Therefore, the mechanism leading to the unusual penta-diene terminus is still unsolved and will be subject of future investigations. The present study elaborated a basis for further work on siphonazole biosynthesis, especially the unusual last reaction in the biosynthetic process, since methods for the detection of the decaboxylated and dehydrated final product of the two purified proteins (didomain Hyd-ACP of SphH and TE SphJ) have been established.

6 Reference

Ahmad, I., & Aqil, F. (2009). *New Strategies Combating Bacterial Infection*. *New Strategies Combating Bacterial Infection*. http://doi.org/10.1002/9783527622931

Alt, S., Burkard, N., Kulik, A., Grond, S., & Heide, L. (2011). An artificial pathway to 3,4-dihydroxybenzoic acid allows generation of new aminocoumarin antibiotic recognized by catechol transporters of E. coli. *Chemistry and Biology*, *18*(3), 304–313. http://doi.org/10.1016/j.chembiol.2010.12.016

Arnez, J. G., & Moras, D. (1997). Structural and functional considerations of the aminoacylation reaction. *Trends in Biochemical Sciences*. http://doi.org/10.1016/S0968-0004(97)01052-9

August, P. R., Tang, L., Yoon, Y. J., Ning, S., Müller, R., Yu, T. W., … Floss, H. G. (1998). Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of Amycolatopsis mediterranei S699. *Chemistry & Biology*, *5*(2), 69–79. http://doi.org/10.1016/S1074-5521(98)90141-7

Auldridge, M. E., Guo, Y., Austin, M. B., Ramsey, J., Fridman, E., Pichersky, E., & Noel, J. P. (2012). Emergent Decarboxylase Activity and Attenuation of / -Hydrolase Activity during the Evolution of Methylketone Biosynthesis in Tomato. *The Plant Cell*. http://doi.org/10.1105/tpc.111.093997

Baker, A., & Cotten, M. (1997). Delivery of bacterial artificial chromosomes into mammalian cells with psoralen-inactivated adenovirus carrier. *Nucleic Acids Research*, *25*(10), 1950–1956. http://doi.org/10.1093/nar/25.10.1950

Belogurov, G. A., Vassylyeva, M. N., Sevostyanova, A., Appleman, J. R., Xiang, A. X., Lira, R., … Vassylyev, D. G. (2009). Transcription inactivation through local refolding of the RNA polymerase structure. *Nature*, *457*(7227), 332–335. http://doi.org/10.1038/nature07510

Benning, M. M., Wesenberg, G., Liu, R., Taylor, K. L., Dunaway-Mariano, D., & Holden, H. M. (1998). The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from Pseudomonas sp. strain CBS-3. *Journal of Biological Chemistry*, *273*(50), 33572–33579. http://doi.org/10.1074/jbc.273.50.33572

Bertani, I., Kojic, M., & Venturi, V. (2001). Regulation of the p-hydroxybenzoic acid hydroxylase gene (pobA) in plant-growth-promoting Pseudomonas putida WCS358. *Microbiology*, *147*(6), 1611–1620.

Birck, M. R., & Woodard, R. W. (2001). Aquifex aeolicus 3-deoxy-D-manno-2-octulosonic acid 8-phosphate synthase: a new class of KDO 8-P synthase? *Journal of Molecular Evolution*, *52*(2), 205–214. http://doi.org/10.1007/s002390010149

Bockarie, M. J., Taylor, M. J., & Gyapong, J. O. (2009). Current practices in the management of lymphatic filariasis. *Expert Review of Anti-Infective Therapy*, *7*(5), 595–605. http://doi.org/10.1586/eri.09.36

Bonanno, J. B., Edo, C., Eswar, N., Pieper, U., Romanowski, M. J., Ilyin, V., … Burley, S. K. (2001). Structural genomics of enzymes involved in sterol/isoprenoid biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 12896–12901. http://doi.org/10.1073/pnas.181466998

Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., & Darst, S. A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell*, *104*(6), 901–912. http://doi.org/10.1016/S0092-8674(01)00286-0

Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., … Gerwick, W. H. (2004). Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium Lyngbya majuscula. *Journal of Natural Products*, *67*(8), 1356–1367. http://doi.org/10.1021/np0499261

Chen, H., & Walsh, C. T. (2001). Coumarin formation in novobiocin biosynthesis: ??-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chemistry and Biology*, *8*(4), 301–312. http://doi.org/10.1016/S1074-5521(01)00009-6

Chen, S., Von Bamberg, D., Hale, V., Breuer, M., Hardt, B., Müller, R., … Leistner, E. (1999). Biosynthesis of ansatrienin (mycotrienin) and naphthomycin: Identification and analysis of two separate biosynthetic gene clusters in Streptomyces collinus Tu 1892. *European Journal of Biochemistry*, *261*(1), 98–107. http://doi.org/10.1046/j.1432-1327.1999.00244.x

Cheng, Y.-Q., Tang, G.-L., & Shen, B. (2003). Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(6), 3149–3154. http://doi.org/10.1073/pnas.0537286100

Clardy, J., Fischbach, M. A., & Walsh, C. T. (2006). New antibiotics from bacterial natural products. *Nature Biotechnology*, *24*(12), 1541–1550. http://doi.org/10.1038/nbt1266

Conti, E., Stachelhaus, T., Marahiel, M. A., & Brick, P. (1997). Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO Journal*, *16*(14), 4174–4183. http://doi.org/10.1093/emboj/16.14.4174

Cosmina, P., Rodriguez, F., de Ferra, F., Grandi, G., Perego, M., Venema, G., & van Sinderen, D. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in Bacillus subtilis. *Molecular Microbiology*, *8*(5), 821–831. http://doi.org/10.1111/j.1365-2958.1993.tb01629.x

Cragg, G. M., & Newman, D. J. (2013). Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta - General Subjects*. http://doi.org/10.1016/j.bbagen.2013.02.008

Criado, L. M., Martín, J. F., & Gil, J. A. (1993). The pab gene of Streptomyces griseus, encoding p-aminobenzoic acid synthase, is located between genes possibly involved in candicidin biosynthesis. *Gene*, *126*(1), 135–139. http://doi.org/10.1016/0378-1119(93)90602-Y

De Baere, I., Derua, R., Janssens, V., Van Hoof, C., Waelkens, E., Merlevede, W., & Goris, J. (1999). Purification of porcine brain protein phosphatase 2A leucine carboxyl methyltransferase and cloning of the human homologue. *Biochemistry*, *38*(50), 16539–16547. http://doi.org/10.1021/bi991646a

Derrer, B., Macheroux, P., & Kappes, B. (2013). The shikimate pathway in apicomplexan parasites: Implications for drug development. *Frontiers in Bioscience : A Journal and Virtual Library*, *18*, 944–69. http://doi.org/10.2741/4155

Dillon, S. C., & Bateman, A. (2004). The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics*, *5*, 109. http://doi.org/10.1186/1471-2105-5-109

Dorrestein, P. C., Blackhall, J., Straight, P. D., Fischbach, M. A., Garneau-Tsodikova, S., Edwards, D. J., … Kelleher, N. L. (2006). Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry*, *45*(6), 1537–1546. http://doi.org/10.1021/bi052333k

Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., … Kelleher, N. L. (2006). Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry*, *45*(42), 12756–12766. http://doi.org/10.1021/bi061169d

Du, L., Chen, M., Sánchez, C., & Shen, B. (2000). An oxidation domain in the BlmIII non-ribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the anti-tumor drug bleomycin in Streptomyces verticillus

ATCC15003. *FEMS Microbiology Letters*, *189*(2), 171–175. http://doi.org/10.1016/S0378-1097(00)00274-3

Duewel, H. S., Radaev, S., Wang, J., Woodard, R. W., & Gatti, D. L. (2001). Substrate and metal complexes of 3-deoxy-D-manno-octulosonate-8-phosphate synthase from Aquifex aeolicus at 1.9-Å Resolution: Implications for the condensation mechanism. *Journal of Biological Chemistry*, *276*(11), 8393–8402. http://doi.org/10.1074/jbc.M007884200

Erol, Ö., Schäberle, T. F., Schmitz, A., Rachid, S., Gurgui, C., El Omari, M., … Könige, G. M. (2010). Biosynthesis of the myxobacterial antibiotic corallopyronin A. *ChemBioChem*, *11*(9), 1253–1265. http://doi.org/10.1002/cbic.201000085

Eys, S., Schwartz, D., Wohlleben, W., & Schinko, E. (2008). Three thioesterases are involved in the biosynthesis of phosphinothricin tripeptide in Streptomyces viridochromogenes Tü494. *Antimicrobial Agents and Chemotherapy*, *52*(5), 1686–1696. http://doi.org/10.1128/AAC.01053-07

Farrar, K., & Donnison, I. S. (2007). Construction and screening of BAC libraries made from Brachypodium genomic DNA. *Nature Protocols*, *2*(7), 1661–1674. http://doi.org/10.1038/nprot.2007.204

Finking, R., & Marahiel, M. A. (2004). Biosynthesis of nonribosomal peptides1. *Annual Review of Microbiology*, *58*, 453–488. http://doi.org/10.1146/annurev.micro.58.030603 .123615

Fisch, K. M. (2013). Biosynthesis of natural products by microbial iterative hybrid PKS–NRPS. *RSC Advances*, *3*(40), 18228. http://doi.org/10.1039/c3ra42661k

Fischbach, M. A., & Walsh, C. T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms. *Chemical Reviews*. http://doi.org/10.1021/cr0503097

Floss, H. G. (1997). Natural products derived from unusual variants of the shikimate pathway. *Natural Product Reports*, *14*(5), 433–452. http://doi.org/10.1039/np9971400433

Floss, H. G., & Yu, T. W. (1999). Lessons from the rifamycin biosynthetic gene cluster. *Current Opinion in Chemical Biology*. http://doi.org/10.1016/S1367-5931(99)00014-9

Furdui, C., Zhou, L., Woodard, R. W., & Anderson, K. S. (2004). Insights into the mechanism of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (Phe) from Escherichia coli using a transient kinetic analysis. *Journal of Biological Chemistry*, *279*(44), 45618–45625. http://doi.org/10.1074/jbc.M404753200

Gehret, J. J., Gu, L., Gerwick, W. H., Wipf, P., Sherman, D. H., & Smith, J. L. (2011). Terminal alkene formation by the thioesterase of curacin A biosynthesis: Structure of a decarboxylating thioesterase. *Journal of Biological Chemistry*, *286*(16), 14445–14454. http://doi.org/10.1074/jbc.M110.214635

Gil, J. A., & Hopwood, D. A. (1983). Cloning and expression of a p-aminobenzoic acid synthetase gene of the candicidin-producing Streptomyces griseus. *Gene*, *25*(1), 119–132.

Gomez-Escribano, J. P., & Bibb, M. J. (2011). Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microbial Biotechnology*, *4*(2), 207–215. http://doi.org/10.1111/j.1751-7915.2010.00219.x

Gosset, G., Bonner, C. A., & Jensen, R. A. (2001). Microbial origin of plant-type 2-keto-3-deoxy-D-arabino-heptulosonate 7-phosphate synthases, exemplified by the chorismate- and tryptophan-regulated enzyme from Xanthomonas campestris. *Journal of Bacteriology*, *183*(13), 4061–4070. http://doi.org/10.1128/JB.183.13.4061-4070.2001

Gu, L., Wang, B., Kulkarni, A., Gehret, J. J., Lloyd, K. R., Gerwick, L., … Sherman, D. H. (2009). Polyketide decarboxylative chain termination preceded by O-sulfonation in curacin A biosynthesis. *Journal of the American Chemical Society*, *131*(44), 16033–16035. http://doi.org/10.1021/ja9071578

Gust B, Chandra G, Jakimowicz D, Yuqing T, Bruton CJ, C. K. (2004). λ Red-Mediated Genetic Manipulation of Antibiotic-Producing Streptomycese. *Adv Appl Microbiol*, *54*(107), 28.

Gust, B., Challis, G. L., Fowler, K., Kieser, T., & Chater, K. F. (2003). PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 1541–1546. http://doi.org/10.1073/pnas.0337542100

Hansen, D. B., Bumpus, S. B., Aron, Z. D., Kelleher, N. L., & Walsh, C. T. (2007). The leading module of mycosubtilin: An adenylation domain with fatty acid selectivity. *Journal of the American Chemical Society*, *129*(20), 6366–6367. http://doi.org/10.1021/ja070890j

Hartmann, M., Heinrich, G., & Braus, G. H. (2001). Regulative fine-tuning of the two novel DAHP isoenzymes aroFp and aroGp of the filamentous fungus Aspergillus nidulans. *Archives of Microbiology*, *175*(2), 112–121. http://doi.org/10.1007/s002030000242

Harvey, A. L., & Harvey, A. L. (2008). Natural products in drug discovery. *Drug Discovery Today*, *13*(19-20), 894–901. http://doi.org/10.1016/j.drudis.2008.07.004

He, J., Magarvey, N., Piraee, M., & Vining, L. C. (2001). The gene cluster for chloramphenicol biosynthesis in Streptomyces venezuelae ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology*, *147*(10), 2817–2829.

Heide, L. (2009). Chapter 18 Aminocoumarins. Mutasynthesis, Chemoenzymatic Synthesis, and Metabolic Engineering. *Methods in Enzymology*. http://doi.org/10.1016/S0076-6879(09)04618-7

Herrmann, K. M., & Weaver, L. M. (1999). THE SHIKIMATE PATHWAY. *Annual Review of Plant Physiology and Plant Molecular Biology*. http://doi.org/10.1146/annurev.arplant.50.1.473

Hertweck, C. (2009). The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition*, *48*(26), 4688–716. http://doi.org/10.1002/anie.200806121

Hill, A. M., Thompson, B. L., Harris, J. P., & Segret, R. (2003). Investigation of the early stages in soraphen A biosynthesis. *Chemical Communications (Cambridge, England)*, (12), 1358–1359. http://doi.org/10.1039/b303542p

Hitchman, T. S., Crosby, J., Byrom, K. J., Cox, R. J., & Simpson, T. J. (1998). Catalytic self-acylation of type II polyketide synthase acyl carrier proteins. *Chemistry & Biology*, *5*(1), 35–47. http://doi.org/10.1016/S1074-5521(98)90085-0

Hoerauf, A., Nissen-Pähle, K., Schmetz, C., Henkle-Dührsen, K., Blaxter, M. L., Büttner, D. W., … Fleischer, B. (1999). Tetracycline therapy targets intracellular bacteria in the filarial nematode Litomosoides sigmodontis and results in filarial infertility. *Journal of Clinical Investigation*, *103*(1), 11–18. http://doi.org/10.1172/JCI4768

Hu, Y., & Floss, H. G. (2001). New type II manumycins produced by Streptomyces nodosus ssp. asukaensis and their biosynthesis. *The Journal of Antibiotics*, *54*(4), 340–348.

Huang, S., Li, N., Zhou, J., & He, J. (2012). [Construction of a new bacterial artificial chromosome (BAC) vector for cloning of large DNA fragments and heterologous expression in Streptomyces]. *Wei Sheng Wu Xue Bao*, *52*(1), 30–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22489457

Irschik, H., Jansen, R., Höfle, G., Gerth, K., & Reichenbach, H. (1985). The corallopyronins, new inhibitors of bacterial RNA synthesis from Myxobacteria. *The Journal of Antibiotics*, *38*(2), 145–152.

Izard, T. (2001). Structural basis for chloramphenicol tolerance in Streptomyces venezuelae by chloramphenicol phosphotransferase activity. *Protein Science : A Publication of the Protein Society*, *10*(8), 1508–1513. http://doi.org/10.1002/pro.101508

Jones, A. C., Gust, B., Kulik, A., Heide, L., Buttner, M. J., & Bibb, M. J. (2013). Phage P1-Derived Artificial Chromosomes Facilitate Heterologous Expression of the FK506 Gene Cluster. *PLoS ONE*, *8*(7). http://doi.org/10.1371/journal.pone.0069319

Jones, A. C., Ottilie, S., Eustáquio, A. S., Edwards, D. J., Gerwick, L., Moore, B. S., & Gerwick, W. H. (2012). Evaluation of Streptomyces coelicolor A3(2) as a heterologous expression host for the cyanobacterial protein kinase C activator lyngbyatoxin A. *FEBS Journal*, *279*(7), 1243–1251. http://doi.org/10.1111/j.1742-4658.2012.08517.x

Jordan, M. A., & Wilson, L. (1998). Microtubules and actin filaments: Dynamic targets for cancer chemotherapy. *Current Opinion in Cell Biology*. http://doi.org/10.1016/S0955-0674(98)80095-1

Kagan, R. M., & Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Archives of Biochemistry and Biophysics*, *310*(2), 417–427. http://doi.org/10.1006/abbi.1994.1187

Kang, Q., Shen, Y., & Bai, L. (2012). Biosynthesis of 3,5-AHBA-derived natural products. *Natural Product Reports*. http://doi.org/10.1039/c2np00019a

Kim, J., & Yi, G.-S. (2012). PKMiner: a database for exploring type II polyketide synthases. *BMC Microbiology*. http://doi.org/10.1186/1471-2180-12-169

Kleinkauf, H., & Von Döhren, H. (1996). A nonribosomal system of peptide biosynthesis. *European Journal of Biochemistry / FEBS*, *236*(2), 335–351. http://doi.org/10.1111/j.1432-1033.1996.00335.x

Knaggs, A. R. (2003). The biosynthesis of shikimate metabolites. *Natural Product Reports*, *20*(1), 119–136. http://doi.org/10.1039/np9951200579

Kohli, R. M., Takagi, J., & Walsh, C. T. (2002). The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(3), 1247–1252. http://doi.org/10.1073/pnas.251668398

König, V., Pfeil, A., Braus, G. H., & Schneider, T. R. (2004). Substrate and metal complexes of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Saccharomyces cerevisiae provide new insights into the catalytic mechanism. *Journal of Molecular Biology*, *337*(3), 675–690. http://doi.org/10.1016/j.jmb.2004.01.055

Konz, D., Klens, A., Schörgendorfer, K., & Marahiel, M. A. (1997). The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chemistry & Biology*, *4*(12), 927–937. http://doi.org/10.1016/S1074-5521(97)90301-X

Kopp, F., & Marahiel, M. A. (2007). Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Natural Product Reports*, *24*(4), 735–749. http://doi.org/10.1039/b613652b

Kourist, R., Jochens, H., Bartsch, S., Kuipers, R., Padhi, S. K., Gall, M., … Bornscheuer, U. T. (2010). The α/β-hydrolase fold 3DM database (ABHDB) as a tool for protein engineering. *ChemBioChem*. http://doi.org/10.1002/cbic.201000213

Kuhstoss, S., & Rao, R. N. (1991). Analysis of the integration function of the streptomycete bacteriophage ???C31. *Journal of Molecular Biology*, *222*(4), 897–908. http://doi.org/10.1016/0022-2836(91)90584-S

Lai, J. R., Koglin, A., & Walsh, C. T. (2006). Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry*. http://doi.org/10.1021/bi061979p

Lee, B. N., Kroken, S., Chou, D. Y. T., Robbertse, B., Yoder, O. C., & Turgeon, B. G. (2005). Functional analysis of all nonribosomal peptide synthetases in Cochliobolus heterostrophus reveals a factor, NPS6, involved in virulence and resistance to oxidative stress. *Eukaryotic Cell*, *4*(3), 545–555. http://doi.org/10.1128/EC.4.3.545-555.2005

Li, C., Hassler, M., & Bugg, T. D. H. (2008). Catalytic promiscuity in the ??/??-hydrolase superfamily: Hydroxamic acid formation, C-C bond formation, ester and thioester hydrolysis in the C-C hydrolase family. *ChemBioChem*, *9*(1), 71–76. http://doi.org/10.1002/cbic.200700428

Li, S., Anzai, Y., Kinoshita, K., Kato, F., & Sherman, D. H. (2009). Functional analysis of MycE and MycF, two O-methyltransferases involved in the biosynthesis of mycinamicin macrolide antibiotics. *ChemBioChem*, *10*(8), 1297–1301. http://doi.org/10.1002/cbic.200900088

Li, W., Ju, J., Rajski, S. R., Osada, H., & Shen, B. (2008). Characterization of the tautomycin biosynthetic gene cluster from Streptomyces spiroverticillatus unveiling new insights into dialkylmaleic anhydride and polyketide biosynthesis. *Journal of Biological Chemistry*, *283*(42), 28607–28617. http://doi.org/10.1074/jbc.M804279200

Liu, H., Jiang, H., Haltli, B., Kulowski, K., Muszynska, E., Feng, X., … He, M. (2009). Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile escherichia coli-Streptomyces artificial chromosome

vector, pSBAC. *Journal of Natural Products*, *72*(3), 389–395. http://doi.org/10.1021/np8006149

Liu, Y. J., Li, P. P., Zhao, K. X., Wang, B. J., Jiang, C. Y., Drake, H. L., & Liu, S. J. (2008). Corynebacterium glutamicum contains 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases that display novel biochemical features. *Applied and Environmental Microbiology*, *74*(17), 5497–5503. http://doi.org/10.1128/AEM.00262-08

Luo, Y., Li, W., Ju, J., Yuan, Q., Peters, N. R., Hoffmann, F. M., … Shen, B. (2010). Functional characterization of TtnD and TtnF, unveiling new insights into tautomycetin biosynthesis. *Journal of the American Chemical Society*, *132*(19), 6663–6671. http://doi.org/10.1021/ja9082446

Ma, N., Wei, L., Fan, Y., & Hua, Q. (2012). Heterologous expression and characterization of soluble recombinant 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase from Actinosynnema pretiosum ssp. auranticum ATCC31565 through co-expression with Chaperones in Escherichia coli. *Protein Expression and Purification*, *82*(2), 263–269. http://doi.org/10.1016/j.pep.2012.01.013

Maeda, H., & Dudareva, N. (2012). The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annual Review of Plant Biology*. http://doi.org/10.1146/annurev-arplant-042811-105439

Magarvey, N. a., Beck, Z. Q., Golakoti, T., Ding, Y., Huber, U., Hemscheidt, T. K., … Sherman, D. H. (2006). Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins. Potent anticancer agents from cyanobionts. *ACS Chemical Biology*, *1*(12), 766–779. http://doi.org/10.1021/cb6004307

Marahiel, M. A. (2009). Working outside the protein-synthesis rules: Insights into non-ribosomal peptide synthesis. *Journal of Peptide Science*. http://doi.org/10.1002/psc.1183

Marahiel, M. A., Stachelhaus, T., & Mootz, H. D. (1997). Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chemical Reviews*, *97*(7), 2651–2674. http://doi.org/10.1016/S0006-3495(02)75175-8

McDaniel, R., Ebert-Khosla, S., Hopwood, D. A., & Khosla, C. (1993). Engineered biosynthesis of novel polyketides. *Science (New York, N.Y.)*, *262*(5139), 1546–1550. http://doi.org/10.1126/science.8248802

Meluzzi, D., Zheng, W. H., Hensler, M., Nizet, V., & Dorrestein, P. C. (2008). Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorganic and Medicinal Chemistry Letters*, *18*(10), 3107–3111. http://doi.org/10.1016/j.bmcl.2007.10.104

Mercer, A. C., & Burkart, M. D. (2007). The ubiquitous carrier protein--a window to metabolite biosynthesis. *Natural Product Reports*, *24*(4), 750–773. http://doi.org/10.1039/b603921a

Mishra, B. B., & Tiwari, V. K. (2011). Natural products: An evolving role in future drug discovery. *European Journal of Medicinal Chemistry*. http://doi.org/10.1016/j.ejmech.2011.07.057

Mizuno, C. M., Kimes, N. E., López-Pérez, M., Ausó, E., Rodriguez-Valera, F., & Ghai, R. (2013). A Hybrid NRPS-PKS Gene Cluster Related to the Bleomycin Family of Antitumor Antibiotics in Alteromonas macleodii Strains. *PLoS ONE*, *8*(9). http://doi.org/10.1371/journal.pone.0076021

Moore, B. S., & Hertweck, C. (2002). Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Natural Product Reports*, *19*(1), 70–99. http://doi.org/10.1039/b003939j

Mootz, H. D., Schwarzer, D., & Marahiel, M. A. (2000). Construction of hybrid peptide synthetases by module and domain fusions. *Proceedings of the National Academy*

of Sciences of the United States of America, *97*(11), 5848–5853. http://doi.org/10.1073/pnas.100075897

Mosberg, J. A., Lajoie, M. J., & Church, G. M. (2010). Lambda red recombineering in Escherichia coli occurs through a fully single-stranded intermediate. *Genetics*, *186*(3), 791–799. http://doi.org/10.1534/genetics.110.120782

Mukhopadhyay, J., Das, K., Ismail, S., Koppstein, D., Jang, M., Hudson, B., … Ebright, R. H. (2008). The RNA Polymerase "Switch Region" Is a Target for Inhibitors. *Cell*, *135*(2), 295–307. http://doi.org/10.1016/j.cell.2008.09.033

Müller, I., & Müller, R. (2006). Biochemical characterization of MelJ and MelK. *The FEBS Journal*, *273*(16), 3768–3778. http://doi.org/10.1111/j.1742-4658.2006.05385.x

Müller, I., Weinig, S., Steinmetz, H., Kunze, B., Veluthoor, S., Mahmud, T., & Müller, R. (2006). A unique mechanism for methyl ester formation via an amide intermediate found in myxobacteria. *ChemBioChem*, *7*(8), 1197–1205. http://doi.org/10.1002/cbic.200600057

Musiol, E. M., Härtner, T., Kulik, A., Moldenhauer, J., Piel, J., Wohlleben, W., & Weber, T. (2011). Supramolecular templating in kirromycin biosynthesis: The acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the trans-AT PKS. *Chemistry and Biology*, *18*(4), 438–444. http://doi.org/10.1016/j.chembiol.2011.02.007

Nadkarni, S. R., Patel, M. V, Chatterjee, S., Vijayakumar, E. K., Desikan, K. R., Blumbach, J., … Limbert, M. (1994). Balhimycin, a new glycopeptide antibiotic produced by Amycolatopsis sp. Y-86,21022. Taxonomy, production, isolation and biological activity. *The Journal of Antibiotics*, *47*(3), 334–341.

Nardini, M., & Dijkstra, B. W. (1999). ??/?? hydrolase fold enzymes: The family keeps growing. *Current Opinion in Structural Biology*. http://doi.org/10.1016/S0959-440X(99)00037-8

Nett, M., & König, G. M. (2007). The chemistry of gliding bacteria. *Natural Product Reports*, *24*(6), 1245–1261. http://doi.org/10.1039/b612668p

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., … Schrag, J. (1992). The alpha/beta hydrolase fold. *Protein Engineering*, *5*(3), 197–211. http://doi.org/10.1093/protein/5.3.197

Paravicini, G., Schmidheini, T., & Braus, G. (1989). Purification and properties of the 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (phenylalanine-inhibitable) of Saccharomyces cerevisiae. *European Journal of Biochemistry*, *186*(1-2), 361–366. http://doi.org/10.1111/j.1432-1033.1989.tb15217.x

Pelzer, S., Süßmuth, R., Heckmann, D., Recktenwald, J., Huber, P., Jung, G., & Wohlleben, W. (1999). Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in Amycolatopsis mediterranei DSM5908. *Antimicrobial Agents and Chemotherapy*, *43*(7), 1565–1573.

Perlova, O., Fu, J., Kuhlmann, S., Krug, D., Stewart, A. F., Zhang, Y., & Müller, R. (2006). Reconstitution of the myxothiazol biosynthetic gene cluster by red/ET recombination and heterologous expression in Myxococcus xanthus. *Applied and Environmental Microbiology*, *72*(12), 7485–7494. http://doi.org/10.1128/AEM.01503-06

Piel, J. (2010). Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports*, *27*(7), 996–1047. http://doi.org/10.1039/b816430b

Podzelinska, K., Latimer, R., Bhattacharya, A., Vining, L. C., Zechel, D. L., & Jia, Z. (2010). Chloramphenicol Biosynthesis: The Structure of CmlS, a Flavin-Dependent

Halogenase Showing a Covalent Flavin-Aspartate Bond. *Journal of Molecular Biology*, *397*(1), 316–331. http://doi.org/10.1016/j.jmb.2010.01.020

Purich, D. L. (2010). *Enzyme Kinetics: Catalysis & Control*. *Enzyme Kinetics: Catalysis & Control*. http://doi.org/10.1016/B978-0-12-380924-7.10005-5

Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., & Walsh, C. T. (1998). Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carder protein domains in peptide synthetases. *Biochemistry*, *37*(6), 1585–1595. http://doi.org/10.1021/bi9719861

Rangaswamy, V., Jiralerspong, S., Parry, R., & Bender, C. L. (1998). Biosynthesis of the Pseudomonas polyketide coronafacic acid requires monofunctional and multifunctional polyketide synthase proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(26), 15469–15474. http://doi.org/10.1073/pnas.95.26.15469

Ray, J. M., Yanofsky, C., & Bauerle, R. (1988). Mutational analysis of the catalytic and feedback sites of the tryptophan-sensitive 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase of Escherichia coli. *Journal of Bacteriology*, *170*(12), 5500–5506.

Rix, U., Fischer, C., Remsing, L. L., & Rohr, J. (2002). Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Natural Product Reports*, *19*(5), 542–580. http://doi.org/10.1039/b103920m

Rouhiainen, L., Paulin, L., Suomalainen, S., Hyytiäinen, H., Buikema, W., Haselkorn, R., & Sivonen, K. (2000). Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in Anabaena strain 90. *Molecular Microbiology*. http://doi.org/10.1046/j.1365-2958.2000.01982.x

Sambrook, J., & W Russell, D. (2001). Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*, 999. Retrieved from

http://books.google.com/books?id=YTxKwWUiBeUC&printsec=frontcover\npapers2://publication/uuid/BBBF5563-6091-40C6-8B14-06ACC3392EBB

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463–5467. http://doi.org/10.1073/pnas.74.12.5463

Sattely, E. S., Fischbach, M. A., & Walsh, C. T. (2008). Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Natural Product Reports*, *25*(4), 757–793. http://doi.org/10.1039/b801747f

Sawitzke, J. A., Thomason, L. C., Costantino, N., Bubunenko, M., Datta, S., & Court, D. L. (2007). Recombineering: In Vivo Genetic Engineering in E. coli, S. enterica, and Beyond. *Methods in Enzymology*. http://doi.org/10.1016/S0076-6879(06)21015-2

Schaberle, T. F., Mir Mohseni, M., Lohr, F., Schmitz, a., & Konig, G. M. (2013). Function of the Loading Module in CorI and of the O-Methyltransferase CorH in Vinyl Carbamate Biosynthesis of the Antibiotic Corallopyronin A. *Antimicrobial Agents and Chemotherapy*, *58*(2), 950–956. http://doi.org/10.1128/AAC.01894-13

Schäberle, T. F., Siba, C., Höver, T., & König, G. M. (2013). An easy-to-perform photometric assay for methyltransferase activity measurements. *Analytical Biochemistry*, *432*(1), 38–40. http://doi.org/10.1016/j.ab.2012.09.026

Schiefer, A., Schmitz, A., Schäberle, T. F., Specht, S., Lämmer, C., Johnston, K. L., … Pfarr, K. (2012). Corallopyronin a specifically targets and depletes essential obligate wolbachia endobacteria from filarial nematodes in vivo. *Journal of Infectious Diseases*, *206*(2), 249–257. http://doi.org/10.1093/infdis/jis341

Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., … Franceschi, F. (2001). Structural basis for the interaction of antibiotics with the

peptidyl transferase centre in eubacteria. *Nature*, *413*(6858), 814–821. http://doi.org/10.1038/35101544

Schnappauf, G., Hartmann, M., Künzler, M., & Braus, G. H. (1998). The two 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase isoenzymes from Saccharomyces cerevisiae show different kinetic modes of inhibition. *Archives of Microbiology*, *169*(6), 517–524. http://doi.org/10.1007/s002030050605

Schofield, L. R., Patchett, M. L., & Parker, E. J. (2004). Expression, purification, and characterization of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase from Pyrococcus furiosus. *Protein Expression and Purification*, *34*(1), 17–27. http://doi.org/10.1016/j.pep.2003.11.008

Schwarzer, D., Finking, R., & Marahiel, M. A. (2003). Nonribosomal peptides: from genes to products. *Natural Product Reports*, *20*(3), 275–287. http://doi.org/10.1039/b111145k

Shawky, R. M., Puk, O., Wietzorrek, A., Pelzer, S., Takano, E., Wohlleben, W., & Stegmann, E. (2007). The border sequence of the balhimycin biosynthesis gene cluster from Amycolatopsis balhimycina contains bbr, encoding a StrR-like pathway-specific regulator. *Journal of Molecular Microbiology and Biotechnology*, *13*(1-3), 76–88. http://doi.org/10.1159/000103599

Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology*. http://doi.org/10.1016/S1367-5931(03)00020-6

Shen, B., Du, L., Sanchez, C., Edwards, D. J., Chen, M., & Murrell, J. M. (2001). The biosynthetic gene cluster for the anticancer drug bleomycin from Streptomyces verticillus ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis. *Journal of Industrial Microbiology & Biotechnology*, *27*(6), 378–385. http://doi.org/10.1038/sj.jim.7000194

Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8794–8797. http://doi.org/10.1073/pnas.89.18.8794

Shumilin, I. A., Kretsinger, R. H., & Bauerle, R. H. (1999). Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from Escherichia coli. *Structure (London, England : 1993)*, *7*(7), 865–875. http://doi.org/10.1016/S0969-2126(99)80109-9

Shumilin, I. A., Zhao, C., Bauerle, R., & Kretsinger, R. H. (2002). Allosteric inhibition of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase alters the coordination of both substrates. *Journal of Molecular Biology*, *320*(5), 1147–1156. http://doi.org/10.1016/S0022-2836(02)00545-4

Siebert, M., Sommer, S., Li, S. M., Wang, Z. X., Severin, K., & Heide, L. (1996). Genetic engineering of plant secondary metabolism. Accumulation of 4-hydroxybenzoate glucosides as a result of the expression of the bacterial ubiC gene in tobacco. *Plant Physiology*, *112*(2), 811–819. http://doi.org/10.1104/pp.112.2.811

Silakowski, B., Kunze, B., & Müller, R. (2001). Multiple hybrid polyketide synthase/non-ribosomal peptide synthetase gene clusters in the myxobacterium Stigmatella aurantiaca. *Gene*, *275*(2), 233–240. http://doi.org/10.1016/S0378-1119(01)00680-1

Silakowski, B., Schairer, H. U., Ehret, H., Kunze, B., Weinig, S., Nordsiek, G., … Müller, R. (1999). New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of Stigmatella aurantiaca DW4/3-1. *Journal of Biological Chemistry*, *274*(52), 37391–37399. http://doi.org/10.1074/jbc.274.52.37391

Sosio, M., Bossi, E., & Donadio, S. (2001). Assembly of large genomic segments in artificial chromosomes by homologous recombination in Escherichia coli. *Nucleic Acids Research*, *29*(7), E37. http://doi.org/10.1093/nar/29.7.e37

Sosio, M., Giusino, F., Cappellano, C., Bossi, E., Puglia, A. M., & Donadio, S. (2000). Artificial chromosomes for antibiotic-producing actinomycetes. *Nature Biotechnology*, *18*(3), 343–345. http://doi.org/10.1038/73810

Stone, N. E., Fan, J. B., Willour, V., Pennacchio, L. A., Warrington, J. A., Hu, A., … Myers, R. M. (1996). Construction of a 750-kb bacterial clone contig and restriction map in the region of human chromosome 21 containing the progressive myoclonus epilepsy gene. *Genome Research*, *6*(3), 218–225. http://doi.org/10.1101/gr.6.3.218

Strieker, M., Tanović, A., & Marahiel, M. A. (2010). Nonribosomal peptide synthetases: Structures and dynamics. *Current Opinion in Structural Biology*. http://doi.org/10.1016/j.sbi.2010.01.009

Subramaniam, P. S., Xie, G., Xia, T., & Jensen, R. A. (1998). Substrate ambiguity of 3-deoxy-D-manno-octulosonate 8-phosphate synthase from Neisseria gonorrhoeae in the context of its membership in a protein family containing a subset of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthases. *Journal of Bacteriology*, *180*(1), 119–127.

Sucipto, H., Wenzel, S. C., & Müller, R. (2013). Exploring Chemical Diversity of α-Pyrone Antibiotics: Molecular Basis of Myxopyronin Biosynthesis. *ChemBioChem*, *14*(13), 1581–1589. http://doi.org/10.1002/cbic.201300289

Swaminathan, S., Ellis, H. M., Waters, L. S., Yu, D., Lee, E. C., Court, D. L., & Sharan, S. K. (2001). Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis*, *29*(1), 14–21. http://doi.org/10.1002/1526-968X(200101)29:1<14::AID-GENE1001>3.0.CO;2-X

Tang, G. L., Cheng, Y. Q., & Shen, B. (2004). Leinamycin Biosynthesis Revealing Unprecedented Architectural Complexity for a Hybrid Polyketide Synthase and Nonribosomal Peptide Synthetase. *Chemistry and Biology*, *11*(1), 33–45. http://doi.org/10.1016/S1074-5521(03)00286-2

Taori, K., Liu, Y., Paul, V. J., & Luesch, H. (2009). Combinatorial strategies by marine cyanobacteria: Symplostatin 4, an antimitotic natural dolastatin 10/15 hybrid that synergizes with the coproduced HDAC inhibitor largazole. *ChemBioChem*, *10*(10), 1634–1639. http://doi.org/10.1002/cbic.200900192

Tapas, S., Kumar Patel, G., Dhindwal, S., & Tomar, S. (2011). In Silico sequence analysis and molecular modeling of the three-dimensional structure of DAHP synthase from Pseudomonas fragi. *Journal of Molecular Modeling*, *17*(4), 621–631. http://doi.org/10.1007/s00894-010-0764-y

Thoden, J. B., Holden, H. M., Zhuang, Z., & Dunaway-Mariano, D. (2002). X-ray crystallographic analyses of inhibitor and substrate complexes of wild-type and mutant 4-hydroxybenzoyl-CoA thioesterase. *Journal of Biological Chemistry*, *277*(30), 27468–27476. http://doi.org/10.1074/jbc.M203904200

Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., & Oppenheim, A. (2007). Recombineering: genetic engineering in bacteria using homologous recombination. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, *Chapter 1*, Unit 1.16. http://doi.org/10.1002/0471142727.mb0116s106

Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S., & Calos, M. P. (2001). Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Molecular and Cellular Biology*, *21*(12), 3926–3934. http://doi.org/10.1128/MCB.21.12.3926-3934.2001

Thykaer, J., Nielsen, J., Wohlleben, W., Weber, T., Gutknecht, M., Lantz, A. E., & Stegmann, E. (2010). Increased glycopeptide production after overexpression of

shikimate pathway genes being part of the balhimycin biosynthetic gene cluster. *Metabolic Engineering*, *12*(5), 455–461. http://doi.org/10.1016/j.ymben.2010.05.001

Tooming-Klunderud, A., Rohrlack, T., Shalchian-Tabrizi, K., Kristensen, T., & Jakobsen, K. S. (2007). Structural analysis of a non-ribosomal halogenated cyclic peptide and its putative operon from Microcystis: Implications for evolution of cyanopeptolins. *Microbiology*, *153*(5), 1382–1393. http://doi.org/10.1099/mic.0.2006/001123-0

Trauger, J. W., Kohli, R. M., Mootz, H. D., Marahiel, M. A., & Walsh, C. T. (2000). Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature*, *407*(6801), 215–218. http://doi.org/10.1038/35025116

Van Lancker, F., Adams, A., Delmulle, B., De Saeger, S., Moretti, A., Van Peteghem, C., & De Kimpe, N. (2008). Use of headspace SPME-GC-MS for the analysis of the volatiles produced by indoor molds grown on different substrates. *Journal of Environmental Monitoring : JEM*, *10*(10), 1127–1133. http://doi.org/10.1039/b808608g

Van Lanen, S. G., Lin, S., & Shen, B. (2008). Biosynthesis of the enediyne antitumor antibiotic C-1027 involves a new branching point in chorismate metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 494–499. http://doi.org/10.1073/pnas.0708750105

Von Döhren, H., Keller, U., Vater, J., & Zocher, R. (1997). Multifunctional Peptide Synthetases. *Chemical Reviews*, *97*(7), 2675–2706. http://doi.org/10.1021/cr9600262

Wagner, T., Shumilin, I. A., Bauerle, R., & Kretsinger, R. H. (2000). Structure of 3-deoxy-d-arabino-heptulosonate-7-phosphate synthase from Escherichia coli: comparison of the Mn(2+)*2-phosphoglycolate and the Pb(2+)*2-phosphoenolpyruvate complexes and implications for catalysis. *Journal of Molecular Biology*, *301*(2), 389–399. http://doi.org/10.1006/jmbi.2000.3957

Walsh, C. T., & Fischbach, M. a. (2010). Natural products version 2.0: Connecting genes to molecules. *Journal of the American Chemical Society*, *132*(8), 2469–2493. http://doi.org/10.1021/ja909118a

Weissman, K. J., & Müller, R. (2008). Protein-protein interactions in multienzyme megasynthetases. *ChemBioChem*. http://doi.org/10.1002/cbic.200700751

Weissman, K. J., & Müller, R. (2010). Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Natural Product Reports*, *27*(9), 1276–1295. http://doi.org/10.1039/c001260m

Welker, M., Dittmann, E., & Von Döhren, H. (2012). Cyanobacteria as a source of natural products. *Methods in Enzymology*, *517*, 23–46. http://doi.org/10.1016/B978-0-12-404634-4.00002-4

Wu, J., Howe, D. L., & Woodard, R. W. (2003). Thermotoga maritima 3-deoxy-d-arabino-heptulosonate 7-phosphate (DAHP) synthase: The ancestral eubacterial DAHP synthase? *Journal of Biological Chemistry*, *278*(30), 27525–27531. http://doi.org/10.1074/jbc.M304631200

Wu, J., & Woodard, R. W. (2003). Escherichia coli YrbI is 3-deoxy-D-manno-octulosonate 8-phosphate phosphatase. *Journal of Biological Chemistry*, *278*(20), 18117–18123. http://doi.org/10.1074/jbc.M301983200

Wu, J., & Woodard, R. W. (2006). New insights into the evolutionary links relating to the 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase subfamilies. *Journal of Biological Chemistry*, *281*(7), 4042–4048. http://doi.org/10.1074/jbc.M512223200

Yu, D., Xu, F., Zeng, J., & Zhan, J. (2012). Type III polyketide synthases in natural product biosynthesis. *IUBMB Life*. http://doi.org/10.1002/iub.1005

Yu, G., Nguyen, T. T. H., Guo, Y., Schauvinhold, I., Auldridge, M. E., Bhuiyan, N., … Pichersky, E. (2010). Enzymatic functions of wild tomato methylketone synthases 1 and 2. *Plant Physiology*, *154*(1), 67–77. http://doi.org/10.1104/pp.110.157073

Yu, T.-W., Bai, L., Clade, D., Hoffmann, D., Toelzer, S., Trinh, K. Q., … Floss, H. G. (2002). The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from Actinosynnema pretiosum. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(12), 7968–7973. http://doi.org/10.1073/pnas.092697199

Zhang, G. Q., Bao, P., Zhang, Y., Deng, A. H., Chen, N., & Wen, T. Y. (2011). Enhancing electro-transformation competency of recalcitrant Bacillus amyloliquefaciens by combining cell-wall weakening and cell-membrane fluidity disturbing. *Analytical Biochemistry*, *409*(1), 130–137. http://doi.org/10.1016/j.ab.2010.10.013

Zhang, Y., Muyrers, J. P., Testa, G., & Stewart, A. F. (2000). DNA cloning by homologous recombination in Escherichia coli. *Nature Biotechnology*, *18*(12), 1314–1317. http://doi.org/10.1038/82449

Zhou, L., Wang, J. Y., Wang, J., Poplawsky, A., Lin, S., Zhu, B., … He, Y. W. (2013). The diffusible factor synthase XanB2 is a bifunctional chorismatase that links the shikimate pathway to ubiquinone and xanthomonadins biosynthetic pathways. *Molecular Microbiology*, *87*(1), 80–93. http://doi.org/10.1111/mmi.12084

7 Appendix

7.1 Primer sequences

Table 7-1. Primers used for the heterologous expression of the complete siphonazole gene cluster

Primer	Sequence (5'–3')
Großer Pcil-R	ACATGTATTCGCCCTTGGATCCC
Kleiner Asel-F	GCATTAATAGGTGACACTATAGAATAC
PGF27 +EC10 lamda-r	TAGAGCAGAATCATCAATGCTTAAGACCGCGAACGG CTCGTCACTATAGGGCGAATTGGG
PGF27 +EC10 lamda-f	GCGGTAGGAATACACAGCGCGGAGCACTCGCTGAT TGGTTCGCCTTCTATCGCCTTCTTG
baej fwd	TAGGATCCCGGCAAAGGACATTGAGC
baej rev	TAGGATCCTGACGCCGACGTAAACAC
Start Up- hom	GCGGCCGCCTGGCCGTCGACATTTAGGTGACACTA TAGAATCGCCAGTCGATTGGCTGAG
Start da	CTTCACCCATGCCCTGCTTC
End down hom	TAATAAATTTGCGGCCGCTAATACGACTCACTATAG GGAGATCGCGGTAGGTCTTACCAAC
End up	TGGCGTACCACCAGAAATTG
Start1 BAC	GGTCGAGCTTGACATTGTAGGACTATATTGCTCTAAT

	AAAGCGAGGCATACCATGAAAGG
Start2 BAC	CGCACTGAGTGGGTGATTTG
End1 BAC	TCTGTATGTACTGTTTTTTGCGATCTGCGTTTCGATC CTAATGCCGGCCTTTGAAT
End2 BAC	TGCAGTTGCCGCATGAACG

Table 7-2. Primers used for heterologous protein expression of domains of the ambigol gene cluster, siphonazole gene cluster and corallopyronin A gene cluster.

Primer	Sequence (5'–3')
DAHP Fa Fwd	GAATTCATGGACATTAACAAGATTCTC
DAHP Fa Rev	AAGCTTAATCTATCCTTTCGTGCC
SphH end fw	CACCTTGTATGGCTATATTCTTGG
SphH end dn	CTAGACAAGCGTTGACATTA
SphJ-topo-Up	CACCATGGCAAGCTTGATCAAG
SphJ-topo-Down	CTATTCGTCAACTATGCTAATC
dn_ACP1_TOPO	CTAGACGAGCCGCAGCGCATAG
up_ACP1_TOPO	CACCATGAGCACGCAGGGGAC

Table 7-3. Primers used for sequencing

Primer	Sequence (5'–3')
T7	TAATACGACTCACTATA
T7 term	CTAGTTATGCTCAGCGGT

7.2 Protein sequences

Protein sequence of Amb7

MHHHHHHDINKILDTNIKSLSVLMAPMQMKEQLPITPVATETVLRGRQAVKEILDGKDS
RKFIIVGPCSIHDVKATLEYAEKLKTLADKVQDKLLILMRVYFEKPRTTIGWKGLINDPDL
DDSFNIQKGLLTARNLLINIAELGLPSATEALDPVTPQYISDLISWAAIGARTIESQTHRE
MASGLSMPVGFKNGTDGNIQVALDAIQSSRNPHHFLGIDQIGQISIFQTKGNVYGHIILR
GGGGQPNFDAATVAWVEKKLENLKLPKRIVIDCSHGNSYKNHQLQTAVFNNVLQQITD
GNQSMIGMMLESNLYEGNQKIPSDLNQLKYGVSVTDKCIGWEETEEIILSAHERLSADR
NVMLHTCGMVVSGTPVRNLMATKG

Protein sequence of SphH Hyd-ACP

HHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTLYGYILGQLAAVLKIDPSRIDVTSELTNY
GVDSLVVTDIHKRFEQDLGSMPVTLLLENTTVAAIANFLQQDYADRVAAFFSPMVAATV
TSQATSNTHEQLIDTADLFALEVGSNQPSSATQPPTAPASITLLRQIEPAAIASELDRYG
DQYAQKLFPAWKQNGGSLVNLTELDANPQLLKHLLVNVADKTQAEVWMIGNGPPLVLI
PAIGLTAPVWINQIQQWAADYRVIVIHQQGYGMTDLTSDISTAAVAKLFISTLDQLGINR
PCHVIGSCFGGVAAQYLTQAYPERVCSLTLCGTFNKNFGLPDIDVSELTIDQMIEGAKM
IGSSINRDFDAVAEGLASDQAQPIVEQARSLLLKSQCVSPLVVMRYITQILTLNGQAWLP
RIQAPTLCLSGNLDTIVAPETSRTISQQIPAGRYIEIPGAGHYPFLTHVDLFEQAVRPFLR
EQEAQLMSTLV

Protein sequence of SphJ thioesterase

MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTMASLIKALAPDPQLPQLFLFPFAGGT
ANSYRGLANALKSHFSVYAIDPQGHINRQEPLLDDLETMVEAYLAALLPLIKPDFTLFGH
SLGGAVVYRLTQRLEQLGHAPVTVFISGYHPPHISDKSTAYLNDDQFLEHIEAMGGVPP
EIGQDQSFMRYFLPIFRADFRATETFIHSDRTKIKAPVFVLNGDKDDDAIKHMQEWSYW
LNQVRYRIFSGPHMYLSSQPELIATYIIECNQAISIVDE

Protein sequence of SphI DAHPS

HHHHHHMIVTIEHAAGFEPIEQVLRCCDEYEASLNLTRDKRYTRIRITVDQAYSELFVRL
KSIVGVAKVETTGKPYPLAAREAKSAASVINVRGIEIGGSDLVVMAGPCSVESREQILAT
AHAVRAAGARILRGGAYKPRTSPYDFRGLGEEGLQLLAEAREATGLAIVTEVMSIADLE
LGAYYADIIQIGARNMQNFSLLHACGSINRPILLKRGPSATIEEWLLAAEYILAAGNEQVI
LCERGIRTYEPFTRNTFDVSAIASIKHLSHLPIIGDPSHGVGKAELVPQMARAAIAAGAD
GLIIEVHPNPAQAWSDGQQSLTFEHFEQLMNDLAPIANALGRQLNPNVQALVVS

Protein sequence of CorI ACP1

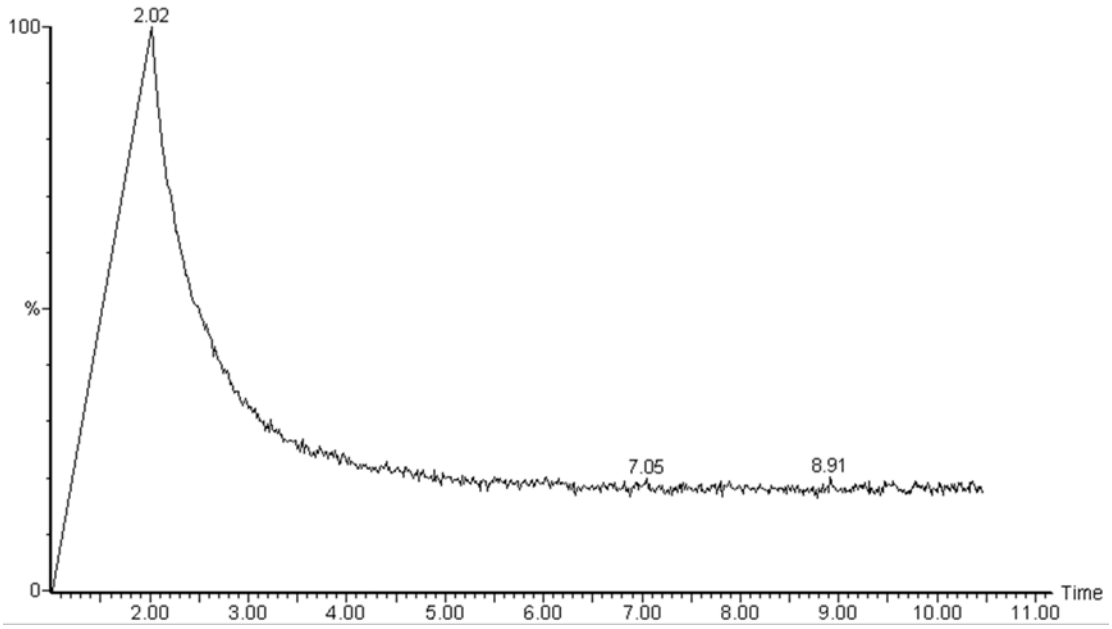
MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTMSTQGTRDVVLVLARRDDLHARLTA
RAKQVEALAGATFVLVKPGRVFRRLGERLFQVDPDDAAHFQRLLPEVEAAHGRVRAIL
HLWNYEAWNVDYAGRRNLKSLTDSIAGNWKFALGAVEKLRGALSGRGPTPLVYVHH
GEGLEAQPHNAIVGAGPAEGAPLVTHPLHVERRLPDLDEASRRLTSALEAVVRGQSLS
GSAAAAAPSAAERLACRKGGVYLVVEPSGAASRPLVEELTGAHGCRVMAVHADALED
FEALRSAVRKRIQDDFGRLDGVIQFLEPSGNIARTLEQVLALDDITSELPLDLFALVAPR
PDAGMAWALARSFNGLRAELTRAGRRLGTSLFVEGPPRLEVREADDAQASAGATRID
FNSEETRAEVLRLLTADVVGIVADMLGKPAASIDLGMSMDLDSASGAHFDSISLTRLAD
RIGTKLGLSLTPILFFKYKKLRAFIAHLHEARYPALSERYALRLV

Protein sequence of CorH *O*-methyltransferase

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEFMSTAYEMVGRTAFVIANWR
AEETEAASPLFSDHVANAFLNAEATEASNAIAAASPSTRYLVRYRTRYFDQTFLDKIAA
GVRQFLILGSGLDTRPIRLAADGVRYFEVEQGHVLEFKRKQLELHGYAQAATFVPAEY
TSVDFLALLEAQGFDFDAQTFILWEGNVFYLEYQSAVNTLIALRDRLRRFEITFDYLSKK
LITKATGYRKSEALLDGFRSLGAPWNTGFDDITELTRTVGLDVAGNFLIADYANDAGLE
MKVDRTLFDDYSIATFANAAGRLDTVAGGRS

7.3 chromatograms of the analyses of the ejected molecule by GC-MS

(A)



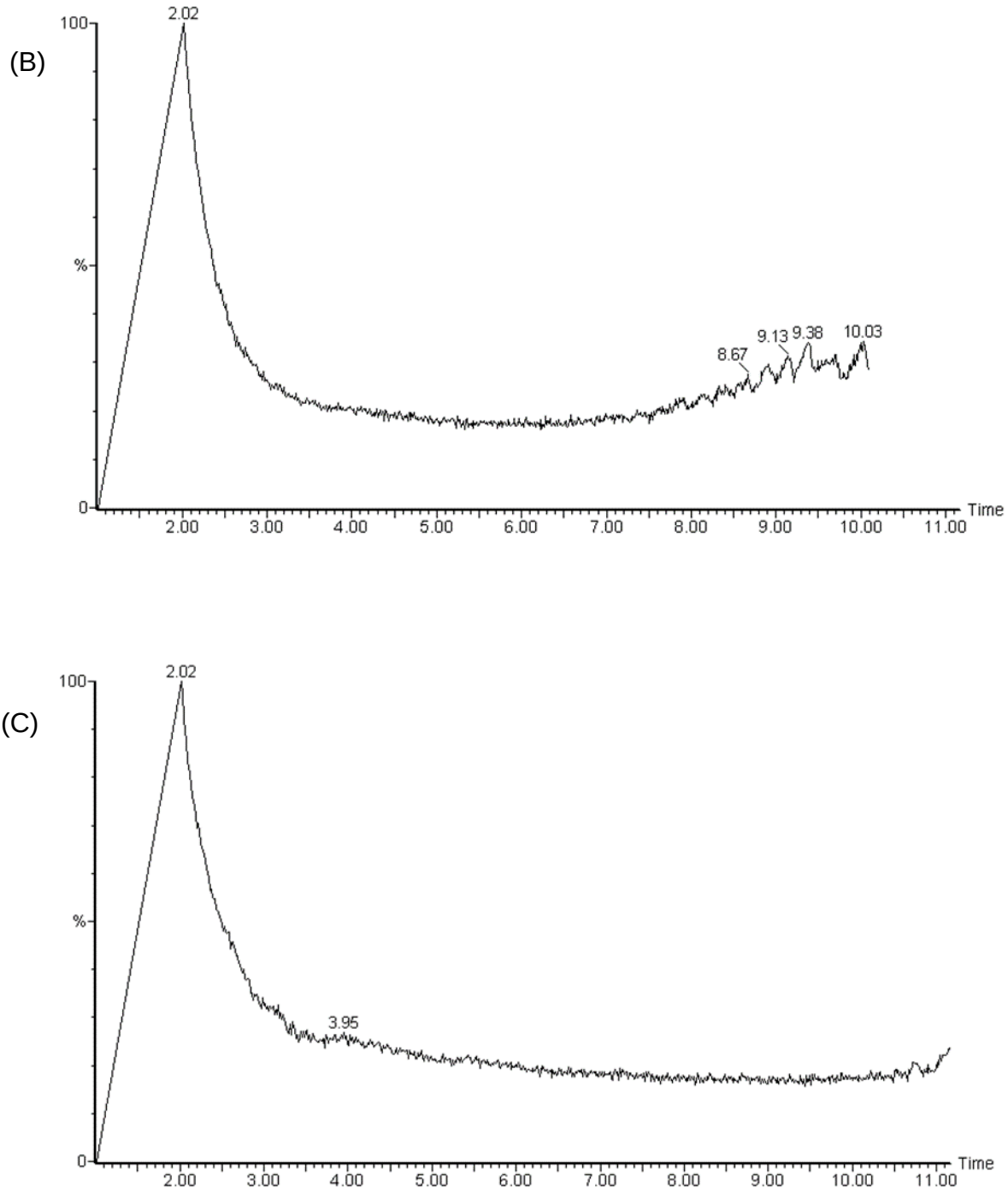


Figure 7-1. GC chromatogram shows no detection of the off-loaded final product in the reaction mixture either during (A) chemical hydrolysis, or by (B) enzymatic process in presence of both proteins, or (C) with a the TE alone.

7.4 Kinetic properties of Amb7

concentration of E4P (µM)	Consumption of PEP (µM/min) first measurment	consumption of PEP (µM/min) second measurment
25	0	0
50	0	10
150	10	20
250	30	50
350	40	70
450	60	80
550	70	80
650	80	80
750	80	80
850	80	80
950	80	80
1000	80	80

Table 7-4. Values applied for calculation of the kinetic properties of Amb7, using Graf Pad Prism 5.

First column: Erythrose-4-phosphate concentrations was varied (25-1000 µM). Column two and three indicate the consumption of PEP, 10 minuets after adding Amb7 at room tempreture of two independent protein purifications, respectively.

7.5 Kinetic properties of SphI

concentration of E4P (µM)	Consumption of PEP (µM/min) first measurment	consumption of PEP (µM/min) second measurment
25	0	0
50	0	0
150	10	20
250	30	30
350	50	30
450	40	40
550	40	40
650	40	40
750	50	40
850	40	40
950	40	40
1000	40	40

Table 7-5. Values applied for calculation of the kinetic properties of Amb7, using Graf Pad Prism 5.

First column: Erythrose-4-phosphate concentrations was varied (0.025-1.000 mM). Column two and three indicate the consumption of PEP 30 minuets after adding SphI at room tempreture of two independent protein purifications, respectively

Publications & Presentations

Schäberle TF, **Mir Mohseni M**, Lohr F, Schmitz A, König GM.(2014). Function of the loading module in CorI and of the O-methyltransferase CorH in vinyl carbamate biosynthesis of the antibiotic corallopyronin A. *Antimicrob Agents Chemother*. 2014; 58(2):950-6.

GM König , S Felder , S Kehraus , S Bouhired , H Harms , T Schäberle , YA Moghaddam, **MM Mohseni**. The secondary metabolome of marine-derived microorganisms. *Planta Med* 2014; 80 - PL2.

Poster Presentations

Mahsa Mir Mohseni, Thomas Höver, Gabriele M. König and Till F. Schäberle BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060. Poster presentation and oral communication at International Workshop of the VAAM Section "Biology of Bacteria Producing Natural Compounds". September 25-26, 2013.Frankfurt.

Oral communications

NRW International Graduate School Biotech-Pharma- Annual Retreat. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY HERPETOSIPHON SP. 060, 10.12.2012, Bonn.

NRW International Graduate School Biotech-Pharma- Annual Retreat. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060, 11.07.2013, Bonn.

NRW International Graduate School Biotech-Pharma- Annual Retreat. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060, 06.05.2014, Bonn.

NRW International Graduate School Biotech-Pharma- Scientific colloquium. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060, 08.11.2013, Bonn.

NRW International Graduate School Biotech-Pharma- Scientific colloquium. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060, 18.01.2013, Bonn.

NRW International Graduate School Biotech-Pharma- Scientific colloquium. BIOSYNTHESIS OF THE ANTIPLASMODIAL NATURAL PRODUCT SIPHONAZOLE BY *HERPETOSIPHON* SP. 060, 14.11.2014, Bonn.